

EMERGENT UNTAUGHT BEHAVIOUR : STIMULUS
EQUIVALENCE AND TRANSITIVE INFERENCE IN
LEARNING DISABLED AND NORMALLY ABLE
PEOPLE

Kerry Teer

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



1997

Full metadata for this item is available in
St Andrews Research Repository
at:

<http://research-repository.st-andrews.ac.uk/>

Please use this identifier to cite or link to this item:

<http://hdl.handle.net/10023/15134>

This item is protected by original copyright

Emergent Untaught Behaviour:
Stimulus Equivalence and Transitive Inference
in Learning Disabled and Normally Able People

KERRY TEER
School of Psychology
University of St. Andrews

Thesis submitted for the degree of
Ph.D.
October 1995



ProQuest Number: 10166198

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10166198

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ABSTRACT

The explanation of the emergence of untaught behaviour has been a topic of considerable interest in behaviour analysis. The experiments in this thesis were designed to examine some of the processes underlying these emergent relations. In doing this, two different paradigms were examined - stimulus equivalence and transitive inference.

The experiments leading to a formal definition of equivalence relations are reported, and the demonstration of cognitive abilities with both humans and non-humans described. The explanations proposed for the basis of stimulus equivalence are discussed. Data from five experiments are then presented. Experiment 1 considered the role of naming in stimulus equivalence and Experiment 2 contrasted this performance with the establishment of transitive inference, both experiments being carried out with adults with learning disabilities as subjects. The results from these experiments suggested that while naming behaviour may help to establish emergent relations, it may not be the basis of stimulus equivalence, and that it may be possible to account for performance on both stimulus equivalence and transitive inference tests in terms of reinforcement contingencies. While subjects who display stimulus equivalence are likely to also display transitive inferences, the reverse relation may not be true.

Experiments 3 and 4 examined the effects of a disruption of the baseline relations on performance on transitive inference and stimulus equivalence tasks. These experiments were both carried out with normally able adults, adults with learning disabilities, and normally developing young children. Experiment 5 was a replication of Experiment 4 with a tighter methodology and a larger number of subjects with learning disabilities. It is suggested that the results obtained in Experiments 4 and 5 can be explained by the development of contextual control of the equivalence relations.

The results from these experiments suggested that the transitive inference and stimulus equivalence paradigms may respectively be concrete and abstract examples of more general emergent relations. These paradigms may also prove to be very useful teaching tools for helping to establish emergent relations.

Declarations

(i) I, Kerry Teer, hereby certify that this thesis, which is approximately 85,397 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

date 17/10/95 signature of candidate

(ii) I was admitted as a research student in October, 1992, and as a candidate for the degree of Doctor of Philosophy in October 1992; the higher study for which this is a record was carried out in the University of St. Andrews between 1992 and 1995.

date 17/10/95 signature of candidate

(iii) I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Philosophy in the University of St. Andrews and that the candidate is qualified to submit this thesis in application for that degree

date 2 October 1995 signature of supervisor

A Unrestricted

In submitting this thesis to the University of St. Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker.

date 17/10/95 signature of candidate

DEDICATION

For my parents, John and Margaret Teer, without whose support this would not have been possible.

ACKNOWLEDGMENTS

Particular thanks are owed to a number of people.

At the University of St. Andrews, special thanks to my supervisor Prof. Chris Cullen for his support and guidance, and to my second supervisor Dr. Rhona Johnston for her assistance.

Also thanks to the staff and residents of Rymonth House in St. Andrews, and in particular to Mr. Ian Davidson (Officer in Charge).

Thanks also to the staff and pupils of the Puffin Playgroup at the School of Psychology, St. Andrews, and to the staff and pupils of the Arellian Nursery School, Belfast for their assistance.

At the University of Leiden, the Netherlands, many thanks to my supervisor Prof. Paul Smeets, and to Dr. Jacqueline Schenk. Thanks also to the staff and pupils of the Elckerlyc Montessori Basisschool, Leiderdorp, and in particular to Mrs. Christine Wallace.

Many, many, thanks to Jill Everett in Psychological and Behavioural Services, Stallington Hall, Stoke-on-Trent, for all her help in collecting the additional data in Experiment 5.

Thanks also to the staff at the Tizard Centre, University of Kent at Canterbury (particularly Peter McGill, Uzzie Laubscher, Lynne Rye, Lisa Sheldon & Julie Brown) and the staff at Psychological Services for People with Learning Disabilities, Coventry Healthcare NHS Trust for support and patience.

Many thanks also to my family and friends for their support throughout my Ph.D.

TABLE OF CONTENTS

CHAPTER	HEADING	PAGE NUMBER
Chapter 1	INTRODUCTION	1
	Stimulus Equivalence	1
	Symbolic Behaviour	2
	Adults with Learning Disabilities	2
	Transitive Inference	3
	Comparison between Stimulus Equivalence and Transitive Inference	4
	Format of the Thesis	4
Chapter 2	THE EARLY EQUIVALENCE EXPERIMENTS	6
	The Basic Findings: Sidman (1971)	6
	Development of the Paradigm	7
	Sidman and Tailby (1982); Sidman, Rauzin, Lazar, Cunningham, Tailby and Carrigan (1982); The definition of Stimulus Equivalence	18
	Summary from early experiments	26
Chapter 3	EXAMPLES OF COGNITION IN NON-HUMANS	29
	The Historical Background	29
	Stimulus Equivalence	33
	Transitive Inference	43
Chapter 4	EXAMPLES OF HUMAN COGNITION	51
	Transitive Inference	51
	Stimulus Equivalence	62
	Sequence Classes	73
	Summary	94
Chapter 5	THE BASIS OF STIMULUS EQUIVALENCE	97
	Naming as the Basis for Stimulus Equivalence	97
	Sidman's Theory of Equivalence	116
	The Relational Frame Account of Equivalence	133

Chapter 6	EXPERIMENT 1: Investigation of the Role of Naming in Equivalence Performance	139
Chapter 7	EXPERIMENT 2: Investigation of Transitive Inference Performance	156
Chapter 8	EXPERIMENT 3: Disruption of Transitive Inference Performance	166
Chapter 9	EXPERIMENT 4: Disruption of Stimulus Equivalence Performance	182
Chapter 10	EXPERIMENT 5: Replication of Experiment 4 - Disruption of Stimulus Equivalence Performance	195
Chapter 11	DISCUSSION	210
	Emergent Relations Require the Establishment of more than the Baseline Relations	210
	Possible Concrete vs. Abstract Relations Distinction	211
	Applications of This Research	214
	General Conclusions	216
	REFERENCES	217
	APPENDIX I	i
	Case Study of Stimulus Equivalence Performance by an Autistic Adult	

CHAPTER 1

Introduction

It is not necessary to learn or be taught everything we know. One of the most important features of human behaviour is the ability to make inferences from previous knowledge, and to integrate new information into well established networks of knowledge. However, it is not clear how these "emergent untaught behaviours" are derived. It may be possible to gain some insight into processes such as learning by examining how these relations are established and maintained.

Stimulus Equivalence

One paradigm which has proved of great utility in examining these emergent relations has been that of stimulus equivalence. This paradigm, which emerged from research by Murray Sidman and his co-workers in the 1970's and early 1980's, provided a framework for examining these relations. Stimulus Equivalence is defined by the emergence of new relations following the establishment of certain conditional discriminations. Sidman based the behavioural paradigm of stimulus equivalence on the mathematical definition of an equivalence relation. The mathematical equivalence relation requires the demonstration of the properties of reflexivity, symmetry and transitivity. Sidman and his co-workers proposed behavioural tests for each of these properties. For example, the reflexivity relation requires the demonstration of generalised identity matching, i.e. given a sample stimulus, A, the subject must be able to match it to an identical comparison stimulus A, without reinforcement or direct training. For the demonstration of the property of symmetry, directly training the subject to select stimulus B when given sample stimulus A, will then result in the subject selecting stimulus A when given sample stimulus B, without direct training or reinforcement, if the original relation established between A and B is a relation of equivalence.

A formal definition and behavioural test for each property was set out in a paper by Sidman and Tailby (1982). This framework provided a way of describing the relations established between stimuli, and a consistent criterion for determining whether the relations established in any experiment were truly emergent equivalence relations. As the tests for each property were behavioural, they allowed for comparisons between different subject populations; normally able adults, subjects with learning disabilities, normally developing young children, and also non-human subjects. Thus Sidman's definition of stimulus equivalence provided a

way to compare emergent relations with different groups of subjects, established under different conditions.

Symbolic Behaviour

Another reason for interest in the behaviours described by stimulus equivalence is that, not only are these behaviours emergent, but they describe "symbolic" relations. The relations established between stimuli in equivalence are not based on formal physical characteristics of the stimuli, but instead are often purely arbitrary. For instance, stimulus equivalence describes the relations between the printed word TWO, the spoken word "TWO" and the numeral 2. These stimuli represent the same concept, but the relations between the stimuli are all arbitrary. There are no physical characteristics that link them, but relations such as these appear to be some of the most durable, and earliest established in normally developing young children. Thus equivalence relations provide a way to examine the emergence of symbolic relations; relations which allow us to tell the time, represent concepts by writing and diagrams, allow us to read maps, and in particular seem relevant to the development of language.

The main point of this thesis then, is to try and look at how relations such as these are established, and to get some idea of the conditions necessary for these emergent relations to be derived. The aim however is not to try and provide a comprehensive explanation of the stimulus equivalence paradigm, (which is too big a task for a single thesis in any case) but rather to try and examine how some of the major research questions apply to performance by a particular group of subjects; adults with learning disabilities.

Adults with Learning Disabilities

There were several reasons for studying equivalence class formation with subjects with learning disabilities. Adults with learning disabilities may well have difficulty acquiring the baseline conditional relations and subsequently deriving equivalence relations, thus making them suitable subjects with which to study how the relations are established. In comparison to normally able adults, these subjects are likely to have simpler behavioural histories and less sophisticated language repertoires, which may make it easier to see which conditions are necessary for the establishment of equivalence relations. Normally developing young children are also likely to have simple behavioural and language repertoires than normally able

adults, but as Saunders and Spradlin (1990) have pointed out, in some ways adults with learning disabilities may make more suitable subjects than young children as any eventual demonstration of equivalence relations is unlikely to be due to developmental changes, or training experiences outside the laboratory.

However, there is another reason for studying the development of equivalence relations with subjects with learning disabilities, "because they are of interest and concern in their own right - they need all the help that is available" (Sidman, 1994, p.533). Research into stimulus equivalence has shown that emergent relations such as these are likely to be a powerful natural learning tool. If subjects with learning disabilities have specific problems with deriving these relations, then understanding the conditions necessary for derived performance may help develop a powerful teaching tool for use with these subjects. A study by Sidman, Kirk & Willson-Morris (1985) had established three six-member equivalence classes; by directly teaching 15 conditional discriminations they were able to demonstrate the emergence of 60 new untaught relations. This demonstrates the potential power of the stimulus equivalence paradigm. While subjects with learning disabilities may have difficulty demonstrating these relation spontaneously, finding ways to support the emergence of these relations could prove very important. This potential has been clear almost from the beginning of research on equivalence. One of the earliest experiments, Sidman (1971), demonstrated the emergence of 40 new relations following direct training on 40 other relations, with an adolescent subject with severe learning disabilities. As Sidman stated, there was now "the prospect of using the technique not only with normally bright children but with people who needed special help to overcome their impaired learning abilities" (Sidman, 1994, p.35).

Transitive Inference

Stimulus equivalence is very much rooted in behaviour analysis, and this thesis mainly examines emergent untaught behaviour from a behaviour analytic perspective. However, behaviour analysis is not the only branch of psychology to have described the emergence of untaught relations. The paradigm of transitive inference has been well established in cognitive and developmental psychology for some time (e.g. Burt, 1919, Piaget, 1928). Transitive inference also requires the derivation of new, unreinforced relations on the basis of explicitly established premises. For instance, given the relational information that stimulus A is larger than stimulus B, and that stimulus B is larger than stimulus C, it is possible to infer

the transitive relation that stimulus A is larger than stimulus C; a transitive inference. Research into transitive inference in the areas of cognitive and developmental psychology provided another description of emergent relations, but established using different procedures. Thus, it seemed possible that studying the conditions necessary for the derivation of transitive inference relations would make a useful comparison to studying the conditions necessary for stimulus equivalence.

Comparison between Stimulus Equivalence and Transitive Inference

Given this potentially useful comparison, it seemed particularly significant that it was possible to demonstrate relations of transitive inference with non-human subjects, while it proved particularly difficult to demonstrate equivalence relations with these subjects. If it was possible to establish why this difference occurred it might indicate the basis for the relations in each paradigm. Thus, several of the experiments in this thesis contrasted performance by the same subjects on both stimulus equivalence and transitive inference. Other experiments contrasted performance by different groups of subjects on the same equivalence/inference tasks.

The main aim of the research in this thesis then, was to examine aspects of the emergence of these untaught relations, and consider some of the conditions which appear to be necessary for these relations to be derived. For instance, two aspects which were considered were, the role of naming behaviour in the development of stimulus equivalence, and the extent to which it is necessary that the baseline stimulus relations form a linear series for the derivation of transitive inference.

The Format of the Thesis

Chapter 2 of the thesis describes some of the early experiments that led to interest in derived relations such as those found in stimulus equivalence. This describes the initial discovery that training certain relations could result in the derivation of new relations, the development of this paradigm, and the formal description of relations of stimulus equivalence as defined by Sidman and Tailby (1982) and Sidman, Rauzin, Lazar, Cunningham, Tailby and Carrigan (1982).

The next two chapters describe the investigation of the paradigms of stimulus equivalence and transitive inference with first non-human, and then human subjects. Chapter 3 describes this investigation with non-human subjects, and also

gives a historical overview of the development of the investigation of cognitive ability with these subjects. Chapter 4 then considers the performance of human subjects on the same paradigms of stimulus equivalence and transitive inference. This chapter also describes some experiments on the development of sequence relations, a topic which seems to connect findings from both equivalence and inference research.

Chapter 5 then examines some of the theories put forward to account for the derivation of equivalence relations. This describes the naming account of equivalence proposed by Dugdale and Lowe (1990) and Horne and Lowe (in press); Murray Sidman's explanation of stimulus equivalence as a fundamental stimulus function; and Steven Hayes' relational frame account of stimulus equivalence.

Chapters 6 to 10 describe the experiments carried out investigating different aspects of stimulus equivalence and transitive inference. These include studies with adults with learning disabilities alone, and also studies comparing performance by normally able adults, adults with learning disabilities, and normally developing young children on the same task. An assessment was also made of the performance of the adults with learning disabilities on a number of independent measures (such as the British Picture Vocabulary Scale), and their performance on these tasks was compared to their performance on the tests of stimulus equivalence and transitive inference.

Chapter 11 provides a general discussion of the results obtained in these different experiments.

Throughout the thesis, the figures and tables for each chapter appear together, following the text, at the end of each chapter.

CHAPTER 2

The Early Equivalence Experiments: 1971-1982

The formal stimulus equivalence paradigm as defined by Sidman and Tailby (1982) developed from work carried out in the 1970's. It arose from a need to explain the results of conditional discrimination training with adolescents with learning disabilities. It was found that this training led to the emergence of more stimulus-stimulus relations than had been established by direct training.

The early papers on these findings discussed the results in terms of cross-modal transfer. For example, a subject might be able to relate auditory words to pictures. In the experiment that subject would then be taught to relate auditory words to visual words. These two auditory-visual equivalences then led to the development of purely visual word-picture equivalences, although these relations had never been directly taught.

The basic methodology and finding were described by Sidman (1971), who considered the skills involved in reading in terms of auditory-visual equivalences.

The Basic Findings: Sidman (1971)

Reading was considered as a type of stimulus - response relation in which the controlling stimuli are visual words. Within that general type of stimulus-response relation several sub-categories were identified, e.g.:

Oral-reading - if a child is shown the visual word *boy* and the subject then says "*boy*", this shows that the child can read the word orally.

Auditory-receptive reading - in this case if the word "*boy*" is said, then the subject should be able to select the word *boy* from several other printed words.

Reading-comprehension - similarly, if the subject is shown the printed word *boy* and is then able to select a picture of a boy from several other pictures, this indicates that the subject understood the word.

The subject in this experiment was an adolescent boy with severe learning disabilities. Prior to training the subject showed good auditory comprehension and picture naming (given spoken names the subject was able to select appropriate pictures, colours or numbers; when shown pictures the subject could name them, but not the corresponding printed words). The subject showed little or no reading comprehension or oral reading (could not match the visual names to the appropriate pictures and could not name the printed words).

During the experiment, the subject was taught to match spoken words to printed words (auditory-receptive reading). Tests then evaluated the effects of this direct teaching on performance on reading comprehension and word naming (see Fig. 2:1). Following this training the subject was able to match printed word samples to picture comparisons and vice-versa, and was also able to name the printed words. Previously the subject had not demonstrated any of these equivalences (Fig.2:1 Equivalences III, IV or VI.)

These results were discussed in terms of the cross-over from auditory-visual equivalences to purely visual equivalences, or cross-modal transfer.

Development of the Paradigm

A series of papers through the 1970's replicated and extended these findings and attempted to quantify and describe further the processes involved.

Sidman and Cresson (1973)

Sidman and Cresson (1973) considered the possibility that "some retarded children who have not achieved the transfer from auditory to visual comprehension have the capability of doing so, but have simply not been taught effectively" (Sidman & Cresson, 1973, p.516). This paper replicated Sidman (1971) with additional control procedures and with two children who had more severe learning disabilities than the original subject. The subjects were first taught identity matching and auditory comprehension. They were then taught to match dictated words to printed words. This resulted in the emergence of reading comprehension in the same manner as was found in Sidman (1971) (see Fig.2:1).

Both Sidman (1971) and Sidman and Cresson (1973) demonstrated mediated transfer, using match-to-sample procedures with subjects with learning disabilities. These papers were concerned with, and considered their results in terms of, the principles of teaching elementary reading. A study by Sidman, Cresson and Willson-Morris (1974) directly examined the nature of the emergent relations between the stimuli.

Sidman, Cresson and Willson-Morris (1974)

In the original findings both sets of visual stimuli were related to the same set of auditory dictated stimuli. The emergent ability to match printed words to pictures might have been mediated by the dictated words - "receptive" mediation (see left half of Fig. 2:2). However, another possible explanation might have been

"expressive" mediation. Prior to training the subjects could name the pictures but not the printed words. After learning the cross-modal matching tasks the subjects were then able to name the printed words. The emergent relations between the pictures and the printed words might therefore have been mediated by the spoken words (see right half of Fig. 2:2). The purpose of this experiment was to determine if the pre-requisites for the emergence of B-C and C-B were B-A and C-A or B-D and C-D (see Fig. 2:2).

In this experiment the authors tried to eliminate the possibility of expressive mediation so as to assess the possibility of receptive naming. To do this, the training relations established were different to those used in Sidman (1971) and Sidman and Cresson (1973). To prevent any direct association between printed words and auditory names, the subject was taught B-A and C-B and tested for the emergence of C-A (auditory names and visual names) and C-D (visual names and spoken names).

The data obtained from this procedure suggested that oral naming was not necessary to mediate the emergence of visual-auditory matching. Both subjects showed the emergence of the C-A relations between auditory words and visual words but without the emergence of C-D naming. Sidman et al in fact suggested that oral naming (C-D) might have been produced by the emergence of C-A rather than mediating the emergence of that relation. However the actual basis for the transfer between stimuli was by no means clear and Sidman et al suggested that it was wisest to limit the term mediation to a procedural sense, i.e. that if an association was established between two stimuli and it was not possible to identify the processes intervening, then mediation refers to the observation that the association was brought about by prior learning involving elements other than these two stimuli.

This experiment by Sidman et al (1974) was an attempt to examine how the emergent relations actually come about. It was an attempt to examine what conditions were necessary for this "transfer" itself rather than considering the results in terms of their relation to reading and how reading behaviour develops.

Other researchers, particularly Spradlin and others, also considered the issue of the development of new relations as shown in Sidman (1971) and tried to analyse the conditions necessary for this.

Spradlin, Cotter and Baxley (1973)

Spradlin, Cotter and Baxley (1973) stated that "responding appropriately to new situations on the basis of what has been learned in prior situations is characteristic of human behavior"(p.556), and that this responding "is especially characteristic of complex behaviors involved in language, reading and reasoning" (p.556).

One problem however is how to account for the "transfer" which is necessary for the new behaviour to emerge and this is often discussed in terms of "concepts". Spradlin et al suggested that it was helpful to consider concepts in terms of stimulus classes such as those defined by Goldiamond (1962). In Goldiamond's definition a stimulus class is a set of stimuli which control similar responses and this definition does not require similar physical characteristics among the members of the stimulus class. Spradlin et al suggested that stimulus classes such as these would provide a mechanism for the transfer in human behaviour and that the term "stimulus class" was likely to create less confusion than the term "concept". However, Spradlin et al also noted that other mechanisms might account for transfer. For instance, Jenkins and Palermo's (1964) response equivalence paradigm was applicable to Sidman's (1971) finding (see Fig. 2:3). Spradlin et al (1973) conducted a series of experiments examining the necessary pre-conditions for the emergence of untaught relations.

The aim of experiment 1 was to investigate if a conditional discrimination could be established without direct training. Two stimuli, A and B, were conditioned to control a common response stimulus C. One of these stimuli, A, was then conditioned to control a new response stimulus D. Probes were then given to see if the second conditional stimulus B also controlled the new response D. The subjects were first given pre-training on the relation A-B. Training then established the relations A-C, B-C and A-D. Probes were then given to assess if the B-D conditional discrimination would be established without direct training. All three subjects with learning disabilities responded appropriately on the B-D probes. For two of the subjects the contingencies for the A-C, B-C and A-D training pairs were then reversed. Both subjects reversed their responses for both the training trials and the probe trials, thus ruling out the possibility that extra-experimental factors might have been responsible for the apparent transfer demonstrated.

Experiment 1 had shown that a conditional discrimination could be developed without direct training. Experiment 2 was very similar to experiment 1 but was designed to examine if the A-B pre-training given in experiment 1 was necessary

for transfer to occur. Three different subjects with learning disabilities were taught the relations A-C, B-C and A-D and performance was then assessed on the test relation B-D. All three subjects showed a very high proportion of criterion responses on the B-D probes. Again, when the contingencies for the training relations were reversed, all three subjects reversed their responses to both the training pairs and the test probes. This experiment suggested that the A-B pre-training given in experiment 1 was not necessary to obtain transfer. Conditioning a common choice response to two conditional stimuli was sufficient to establish similar controlling properties; so that when a new choice response was conditioned to one conditional stimulus the second stimulus also controlled that response.

Experiment 3 investigated the possibility that while A-B training was not necessary to obtain transfer, this training combined with A-D training might be sufficient for B-D transfer to occur. Again, the subjects were three adolescents with learning disabilities. The subjects were given A-B pre-training to 95% or above, accuracy. The subjects were then given A-D training to a similarly high criterion. This time, on the B-D probes only two of the three subjects showed transfer. When the contingencies on the A-D relations were changed, the subjects showed reversal only on the training pairs, not on the probe trials. These results were not conclusive, but as Spradlin et al noted, the procedures in experiment 3 were not ideal for obtaining stable transfer as the A-B slides were never presented after pre-training and this may have adversely affected performance on the probe trials.

Spradlin et al (1973) had conducted a series of experiments trying to replicate the findings of Sidman (1971) and investigate some of the conditions necessary for transfer to occur. Spradlin et al also considered how transfer of this type might be important for learning in general and how failure to exhibit this transfer may be a critical aspect of learning disability.

Dixon and Spradlin (1976)

Dixon and Spradlin (1976) continued this investigation of transfer using stimulus classes. Sidman (1971) had established stimulus equivalence by training the subject to select two different visual stimuli (a picture and a printed word) in response to the same auditory sample stimulus. Following this training the subjects were able to match the two visual stimuli to each other. Similarly, Spradlin et al (1973) established functional equivalence of two different visual stimuli, but using a slightly different procedure. Two visual sample stimuli were conditionally related to one visual comparison stimulus. One of the sample stimuli was then

conditionally related to a new comparison stimulus. Without further training, the second sample stimulus was also found to control responding to the new comparison stimulus.

In an unpublished study Dixon and Spradlin had tried to extend this paradigm by training subjects to select four different visual choice stimuli for each of two visual sample stimuli. The subjects were then taught to conditionally relate each of the samples in response to an auditory stimulus. However, there was no transfer of control from the visual samples to the auditory samples for each class of visual choice stimuli. Dixon and Spradlin (1976) described a series of experiments to examine the conditions necessary for transfer to occur using this paradigm. The overall aim was to determine if a visual class would be established if every member of a set of stimuli was to serve as both sample and comparison stimulus for every other member of the set and, if subjects were trained to select some members of the set in response to an auditory stimulus would they then select all members of the set in response to that stimulus. This was tested in experiment 1. The subjects in all experiments were adolescents with learning disabilities. The visual stimuli were eight abstract symbols divided into two sets, and the auditory stimuli were two nonsense syllables.

In experiment 1 the subjects were trained to classify the eight visual stimuli into two classes using a series of match-to-sample tasks in which each stimulus was presented as both a sample and a comparison stimulus for other members of that class. After the visual discrimination training the subjects were given auditory discrimination training. In the auditory training the subjects were taught to match one of the visual stimuli from each class to one of the nonsense syllables. The subjects were then given auditory probes to see if they selected the visual stimulus that was in the same class as the visual stimulus that was correct in the auditory training task.

All six subjects reached criterion on the visual and auditory training tasks but only three subjects showed transfer of auditory control to the untrained visual stimuli. This failure could have been caused by the visual-matching training being insufficient to establish a stimulus class or alternatively the failure could have been related to the auditory stimuli. This was examined in experiment 2.

In experiment 2, two of the subjects who failed to show transfer in experiment 1 received training on a new visual conditional discrimination. Two new visual

stimuli served as samples for the two visual stimuli that had been directly related to the auditory stimuli in experiment 1. The subjects then received probe trials to see if this control transferred without further training to the other visual stimuli. Both subjects eventually responded correctly to over 90% of the visual probe trials. Following the probes for visual transfer the subjects again received auditory probes to see if they would show transfer to the auditory stimuli following correct responding with the new visual stimuli. One subject did demonstrate this transfer of auditory control while the other did not.

Although both subjects had failed to show transfer of auditory control in experiment 1, both subjects showed transfer of control to the new visual stimuli in experiment 2. This suggested that the previous visual discrimination training had established two visual stimulus classes rather than a series of independent visual discriminations. As the new visual sample stimuli had the same relation as the auditory sample stimuli to the visual choice stimuli, this suggested that the failure to demonstrate transfer may have been linked to the auditory sample stimuli.

Dixon and Spradlin suggested that this failure may have been due to the auditory labels being related to only one member of each class. They suggested that this may have caused the subject to respond to that label as a name for a specific stimulus rather than as a class name. Possibly relating the stimuli to two or more members of the class would cause the label to be extended to all members of the class. This was tested in experiment 3.

The subjects in experiment 3 were the remaining subject who failed to demonstrate transfer in experiment 1 and the subject who failed to demonstrate transfer of auditory control in either experiment 1 or experiment 2. The subjects were given auditory training as in experiment 1 but with a new pair of visual stimuli as comparisons for the same auditory samples. One subject showed transfer following this training. The other subjects performance improved but remained below 90% correct. This subject received auditory training with a third set of stimuli and demonstrated transfer following this. This experiment demonstrated that direct training given to only some members of a class could transfer to the other members of the class without further training. However, it was also clear that the ease with which this transfer occurred varied considerably across subjects. The aim of experiment 4 was to determine the extent of this transfer with one of the subjects who had demonstrated transfer in experiment 1.

In experiment 4, the subject was trained to match two new visual stimuli to the visual stimuli which had been matched to the auditory stimuli in experiment 1. Probes showed that the auditory control established in experiment 1 transferred to these new stimuli. The subject was then trained to relate two other new stimuli to the auditory samples used in experiment 1. The subject was then given probe trials where the two new stimuli were samples and the rest of the visual stimuli were comparisons. The subject selected the visual stimuli in accordance with the classes established. This experiment demonstrated that once a stimulus class had been established it was possible to add new members to the class with relatively little training. The new members of the class would also have all the controlling properties established for the original members of the class. It was also possible to establish new controlling properties with relatively little training.

This series of experiments investigated the conditions under which a stimulus class could be established and how that class could be brought under auditory control by relating an auditory label to only some members of the class. These experiments were intended to investigate the ways in which human subjects are able to respond appropriately in new situations on the basis of previously learned information. As Dixon and Spradlin pointed out, this behaviour occurs all the time in a "natural" setting but had been difficult to investigate in a laboratory setting. The manipulations carried out in this study seemed to provide a way of examining the conditions necessary for transfer of this sort to occur.

Spradlin and Dixon (1976)

A paper by Spradlin and Dixon (1976) continued this analysis of stimulus classes and labels in the development of untaught behaviour. It was similar in nature to that of Dixon and Spradlin (1976) and provided some support for the hypotheses suggested in that paper. Spradlin and Dixon (1976) further refined the definition of a stimulus class as a concept. Spradlin et al (1973) had suggested that it was helpful to regard a concept as similar to Goldiamond's (1962) definition of a stimulus class; a set of stimuli which control common responses. Spradlin and Dixon felt that even this definition may be misleading. They pointed out that, in a typical match-to-sample task, there is no topographical difference between correct and incorrect responses. Rather, it is the relationship between the stimuli that determines if a response is correct or not. Spradlin and Dixon thus defined a concept as "stimuli or events which are substitutable for each other within a given context" (p.555) and if two or more stimuli are established as functionally equivalent in one condition it is likely that they will be functionally equivalent in a

second condition, even without direct training in that situation. Spradlin and Dixon attempted to expand the paradigm used by Spradlin et al for establishing a conditional discrimination without direct training. Spradlin et al had suggested that if several choice stimuli were brought under the control of a single sample stimulus, and a verbal label was then conditioned to control responses to one of the choice stimuli, then this control should transfer to the other choice stimuli. However, when Spradlin and Dixon tested this they found that the auditory labels did not control responding to the other choice stimuli. The aim of Spradlin and Dixon then was to see if there were any non-verbal procedures which could be used to establish these stimulus classes and produce transfer of this sort. Research had indicated that there are a number of ways of establishing associations among stimuli. Jenkins and Palermo (1964) reviewed these and Spradlin and Dixon summarised them as:

Stimulus Equivalence Paradigm: if A controls B and C controls B, then A tends to control C.

Response Equivalence Paradigm: if A controls B and A controls C, then B tends to control C.

Mediation Paradigm: if A controls B and B controls C, then A tends to control C.

Four-Stage Paradigm: if A controls B and C controls B and A controls D then C tends to control D.

Spradlin and Dixon (1976) hypothesised that if each stimulus in a set was used as both sample and comparison for the other members in a visual discrimination task, when an auditory label is conditionally related to one member of a set, the processes involved in all these paradigms should operate on all the remaining members of a set and produce the maximum opportunity for transfer to occur. Spradlin and Dixon tested this possibility in this experiment.

The subjects in this experiment were two adolescents with learning disabilities. Match-to-sample training procedures were used to establish two classes of visual stimuli. Each stimulus within a set served as a sample and a comparison stimulus for the other members within that set. The subjects were then given auditory discrimination training. One visual stimulus from each class was related to one of two nonsense syllables. The remaining three stimuli from each class were used as probes to see if the subjects would select a visual stimulus that was in the same class as the visual stimulus directly related to each of the nonsense syllables. Both subjects reached criterion on the visual discrimination and auditory training trials. However, when given the probe trials, neither subject showed much evidence of

class transfer. The subjects therefore received auditory training with another visual stimulus from each class and the probe trials were given once more. This time both subjects showed transfer of control, responding at above 90% class consistent responses. When shown the visual stimuli and asked to name them, both subjects provided the correct name for a very high proportion of the stimuli.

Thus, the experiment showed that, when every stimulus in a set functioned as both a sample and a comparison for other stimuli in that set, and at least two stimuli from that set are given a label, that label will also be applied to the remaining stimuli in that set. This seemed consistent with Spradlin and Dixon's stated hypothesis that stimuli established as equivalent by direct training in some situations will have similar functions in different situations without direct training. The fact that transfer of function did not seem to occur until the auditory label had been applied to at least two stimuli also fits with Dixon and Spradlin's (1976) suggestion that a label may initially be treated as a name for a specific stimulus until it is related to two or more stimuli, when it may come to function as a class name.

Spradlin and Dixon also noted that it is possible for transfer to occur when stimuli are given only one common function (as in Sidman 1971, Sidman and Cresson 1973, Sidman et al 1974, Spradlin et al 1973). However the results from Spradlin and Dixon (1976) and Dixon and Spradlin (1976) suggest that if stimuli are given equivalent functions in several contexts, then it is more likely that this function will transfer without additional training to a new context. Extensive training such as this may prove to be especially important for young children or people with learning disabilities.

Van Biervliet (1977)

VanBiervliet (1977) used the procedures developed in the studies up to this point and applied them to trying to establish word-object associations with subjects with learning disabilities. VanBiervliet's rationale was that the association of words with objects is thought to be crucial for language acquisition, but establishing these associations with some learning disabled individuals could be extremely difficult using standard training techniques. An alternative technique of pairing manual signs with words and objects had been suggested (Bricker 1972) but there was some resistance to the use of signs on the basis that children taught to communicate manually might not learn to speak but instead communicate only with their hands. Bricker's research had provided some evidence that sign-word and sign-object

training facilitated the acquisition of word object associations but had investigated only receptive word-object associations, not how word-object production might be affected.

VanBiervliet's study then had two purposes: firstly, to determine if sign-object and sign-word training facilitated both receptive and productive word-object associations. In doing this the second purpose was to test Spradlin and Dixon's (1976) functional equivalence proposal; i.e. that if two or more stimuli are established as functionally equivalent by training, there is an increased probability that they will be functionally equivalent in a second condition even without direct reinforcement.

The subjects were 6 retarded adolescents. The stimuli were five nonsense words, 5 junk objects and five nonsense manual signs. The subjects were taught: 1) Object matching; 2) Sign imitation; 3) Object-sign production; 4) Sign-object reception, 5) Word imitation; 6) Word-sign production; 7) Sign-word production. (See Fig. 2:4). This training established the relations between each stimulus and itself (relations 1, 2 and 5) - e.g. in object matching (relation 1) when shown an object as the sample stimulus, the subject had to select the same object from the comparison stimuli. The training also established the relations between objects and signs (relations 3 and 4) and words and signs (relations 6 and 7). For example, in object sign production (relation 3) when given an object as a sample the subject had to produce the appropriate sign. Sign-object reception (relation 4) was the reverse of this, when shown a sign as sample the subject had to select the appropriate object. When the subjects had reached criterion on these tasks they were probed for the development of associations between the words and the object; word-object reception (relation 8) and object-word production (relation 9). All the subjects successfully completed the training tasks. On the probe tasks all subjects showed appropriate associations between words and objects.

Thus, following sign-object and sign-word training, all the subjects were able to correctly associate words and objects, which suggested that this might indeed be a viable way of training spoken language for some individuals. While this sort of training might not be necessary for normally developing children it might be more effective than normal teaching procedures for children with learning disabilities. VanBiervliet suggested that Spradlin and Dixon's (1976) functional equivalence proposal provided an explanation for the acquisition of these word-object associations. If two stimuli control the same response, they are likely to have the

same function in other behavioural contexts, i.e. the sign production was controlled by both the object and the dictated word and so the object and the word became functionally equivalent.

Lazar (1977)

Lazar (1977) used another procedure to investigate the development of functional equivalence classes. A sequential-response training paradigm was used to establish two stimulus classes - "first" and "second". The stimuli used in this procedure were then related to a new set of stimuli by means of a match-to-sample procedure. Finally, probes were given to see if the new stimuli had acquired the functions of "first" and "second" through equivalence in the matching to sample.

The subjects were three normally able adults. Four pairs of stimuli were presented and the subjects were taught always to point to one stimulus in each pair first and the other stimulus second, regardless of spatial position. The aim of this was to establish four baseline sequences/pairs $A1 \rightarrow A2$, $B1 \rightarrow B2$, $C1 \rightarrow C2$, $D1 \rightarrow D2$. Probes were then given to see if this training had established two classes: one consisting of stimuli to which the subjects responded first and one consisting of stimuli to which the subjects responded second. The probes consisted of presentations of "cross-over" pairs; for example, given the test pair $A2/B1$, the subjects were expected to respond to stimulus $B1$ first followed by $A2$. This would suggest that the subjects behaviour was controlled by the classes "first" and "second". These original stimuli were then related to two new pairs of stimuli, $E1$ and $E2$, and $F1$ and $F2$, using a match-to-sample procedure. The original stimuli served as samples and the new stimuli as comparisons, so that, for example, when the sample was stimulus $A1$, stimulus $E1$ was the correct comparison and when the sample was $A2$, $E2$ was the correct comparison. Following this training, probes were again given to see if the new stimuli $E1$, $E2$, $F1$ and $F2$ had acquired the functions of "first" and "second" as a result of being conditionally related to the original A , B , C and D stimuli.

All three subjects learned the four baseline sequences, and probes demonstrated the establishment of the two classes "first" and "second". All three subjects learned the matching-to-sample tasks. Following this training, two subjects immediately responded to the stimuli in predictable sequences, pointing to $E1$ and $F1$ first and $E2$ and $F2$ second. The third subjects performance showed no evidence of transfer of function having occurred. Subsequent tests showed that the sample and comparison stimuli in the matching-to -sample tasks had not become functionally

equivalent. This subject appeared to have treated the matching-to-sample trials as a number of unrelated tasks and, as a result, the transfer of sequence function would not have been expected.

The baseline training in the first part of this experiment established two sequence classes - "first" and "second" for all three subjects. When matching-to-sample was used to relate new stimuli to the stimuli in these sequence classes, two of the three subjects responded to these new stimuli in predictable sequences without further training. Thus, stimulus class membership can be extended by match-to-sample procedures, even when the original stimulus classes were not established in a match-to-sample context.

Lazar (1977) suggested that studies of this sort might be relevant to the study of language. Much of this behavioural repertoire seems to emerge without direct training and had been difficult to account for using a functional analysis. Deese (1965) and Jenkins (1965) had suggested that the grammatical elements could be viewed as members of syntactical classes. Braine (1963) and Browne (1973) had shown some evidence that an early indication of syntactical structure was word order. Lazar suggested that words occupying equivalent positions might be members of the same grammatical class and so the results from this study might be similar to a child learning to combine an adjective and a noun appropriately. This study was "an initial attempt to apply a functional analysis to simple sequences that may be analogous to syntactical structures" (p.392).

Sidman & Tailby (1982);

Sidman, Rauzin Lazar, Cunningham, Tailby & Carrigan (1982):

The Definition of Stimulus Equivalence

Sidman and Tailby (1982) and Sidman, Rauzin, Lazar, Cunningham, Tailby and Carrigan (1982) carefully analysed the procedures used in trying to establish equivalence relations.

Sidman and Tailby (1982)

Sidman and Tailby (1982) considered the precise status of the relations between the stimuli in equivalence classes, and proposed a definitive framework of relations and behavioural tests for the consistent demonstration of equivalence relations. This definition proposed precise tests for determining if relations established between stimuli were conditional relations, or equivalent and could be called

matching-to-sample. For example, in a conditional discrimination a subject learns to select stimulus B1 in the presence of stimulus A1 and stimulus B2 in the presence of A2. Sidman and Tailby stated that the relation between these stimuli, "if . . . then", was directly observable by reference to the subjects ongoing interaction with the procedure. However, a well established conditional discrimination may demonstrate not only the "if . . . then" conditional relations between the stimuli, but also an equivalence relation. If this equivalence relation is assumed the subjects performance may often be referred to as "matching-to-sample". However, to determine if a conditional relation is also an equivalence relation requires additional tests. Sidman and Tailby specified the tests necessary to demonstrate equivalence. These were derived from the mathematical definition of an equivalence relation and specified three properties: reflexivity, symmetry and transitivity.

Reflexivity

If a conditional relation is reflexive each stimulus must bear that relation (r) to itself; if A then A - ArA. The demonstration of reflexivity requires generalised identity matching.

Symmetry

The conditional relation between two stimuli must also be reversible if the stimuli are equivalent. This requires the demonstration of symmetry, so that if two stimuli are conditionally related - A-B, it should then be possible to demonstrate B-A without further training. The demonstration of symmetry then is functional sample-comparison reversibility.

Transitivity

The demonstration of the third property, transitivity, requires three stimuli. If two conditional relations, A-B and B-C are established, then transitivity requires the demonstration of A-C. The transitivity relation shows the relation between the sample of the first conditional discrimination and the comparison of the second.

Sidman and Tailby (1982) stated that in order for a conditional relation to be called "matching-to-sample" all three properties of reflexivity, symmetry and transitivity must be demonstrated. Sidman and Tailby also detailed combined tests for evaluating symmetry and transitivity simultaneously. For example, if a subject was directly taught the relations A-B and A-C, the demonstration of B-C and C-B would require the properties of both symmetry and transitivity. Symmetry of the

relations A-B and A-C would give B-A and C-A. Transitivity of B-A and A-C would give B-C while transitivity of C-A and A-B would give the relation C-B.

Sidman and Tailby then described an experiment which replicated the procedures of Sidman (1971), Sidman and Cresson (1973) and Spradlin, Cotter and Baxley (1973). The results from these experiments were then related to Sidman and Tailby's definition of equivalence, and this training and test paradigm was also extended to consider 4-member equivalence classes. In Sidman (1971) and Sidman and Cresson (1973) the subjects learned or demonstrated the conditional discrimination AB (dictated names \rightarrow pictures). They were then taught the conditional discrimination AC (dictated names \rightarrow printed names). The subjects then demonstrated that these conditional relations were also equivalence relations by demonstrating the relations BC and CB (matching pictures and printed names and vice versa). These relations required a combination of symmetry and transitivity, the subjects were not explicitly trained to demonstrate these relations, and had not been able to relate these pictures and printed words to each other prior to learning the AB and AC conditional discriminations. The subjects also proved able to name the pictures (BD) and the printed words (CD) aloud (see fig. 2:1)

One of the main purposes of the experiment in Sidman and Tailby (1982) was to add one further stimulus to each class and so investigate the power of equivalence relations to generate a network of interchangeable stimuli (see Fig. 2:5). The ABC relations in this experiment, established by training AB and AC, represented a replication of the results from Sidman (1971) and Sidman and Cresson (1973). This was then extended by training the relation DC, where new stimuli from set D functioned as samples for stimuli in set C. The aim was to see if these DC relations expanded the three-member ABC class to a four-member ABCD class of equivalent stimuli (see Fig. 2:5). Spradlin, et al (1973) had demonstrated the emergence of relations similar to DB but had not tested for BD. As both DB and the symmetrical relation BD were necessary for the demonstration of 4-member classes, this experiment both replicated and extended Spradlin et al's findings.

The paradigm used in this experiment permitted the testing of equivalence in several ways. Training the relations AB and AC allowed equivalence to be tested by the relations BC and CB. This established the upper triangle in Fig. 2:5. In this case the trained relations shared the same sample stimulus, A. Training a further relation DC established the lower triangle in Fig. 2:5 and equivalence could be tested by the relation AD. In this case the trained relations shared the same

comparison stimulus, C. It was then possible to test for the existence of 4-member classes by testing the relations DB and BD.

The subjects were eight normally able children aged between 5:0 and 7:5 (years:months). As part of a series of pre-tests, the subjects were tested on identity matching with the Greek letters used in the experiment. This served as a test for reflexivity. The subjects were then taught three conditional discriminations. First they were taught AB, then AC, then these two discriminations were presented mixed together. The subjects were then taught DC and these relations were mixed with AB and AC to constitute a mixed baseline. The subjects were then tested for the development of 4-member and 3-member equivalence relations. The probes for each relation were inserted into an appropriate baseline of training relations so that, for example, the DB 4-member equivalence probes were inserted into a baseline of AB, AC and DC relations while the 3-member BC and CB equivalence probes required a baseline of AB and AC relations. Tests were given for both matching-to-sample and oral naming. The matching-to-sample probes were 4-member equivalence relations DB and BD; 3-member equivalence relations AD, BC and CB; and symmetry probes CD. At the end of the experiment the subjects were tested for oral naming of the B, C and D stimuli.

Six of the eight subjects demonstrated the emergence of six new sets of conditional discriminations as a result of being directly taught three sets of conditional discriminations. These performances included the emergence of the 4-member relations as well as all the symmetrical and transitive relations defined as necessary for the relations to emerge. The children also named the stimuli in accordance with the class memberships established by training.

Two subjects failed to demonstrate the emergence of all these new conditional relations. Neither subject demonstrated the transitive relations required for the 4-member DB and BD relations. One of these subjects was given a series of tests to examine the status of the prerequisite relations. This subject's poor performance on the 4-member test was shown to be consistent with the absence of several of the prerequisite relations. Interestingly though, this subject's performance on the naming test was highly accurate. The other subject did not receive all the tests necessary to evaluate the prerequisite relations for equivalence.

These results suggested that Sidman and Tailby's definition of the properties necessary for stimulus equivalence was indeed useful. The six subjects who

demonstrated 3 and 4-member equivalence relations also demonstrated all the relations defined as necessary for these equivalence relations to emerge. For at least one of the subjects who failed to demonstrate 4-member equivalence relations, this failure was predictable because of the failure to establish some of the prerequisite relations.

This experiment also showed how this stimulus equivalence procedure had enormous potential for generating large amounts of new behaviour from only a few directly trained relations. For instance, three conditional discriminations were directly trained, AB, AC and AD but six other conditional relations emerged without direct training, as well as oral naming relations.

Sidman and Tailby suggested that stimulus equivalence was likely to be independent of naming. The children in this experiment were able to name the set D letters with the same class names that had been explicitly taught for the B and C stimuli. This had raised the possibility that naming might be mediating the emergence of the new conditional relations. However, one subject's performance on the naming test suggested that this was unlikely. This subject named the B and C stimuli with no hesitation but was very reluctant to give names for the D stimuli until prompted. Sidman and Tailby suggested that this showed the subject had not named the D stimuli prior to the naming test, which occurred after the test for the emergent conditional relations. Therefore, although naming may facilitate equivalence relations when it occurs, it is not necessary for the development of equivalence. Sidman and Tailby also argued that naming was not sufficient for the development of equivalence. One subject was able to name the B, C and D stimuli in accordance with the predicted equivalence classes, but did not demonstrate the emergence of these classes on the equivalence tests. As Sidman and Tailby pointed out, it is not surprising that naming was not sufficient to establish these equivalence classes as the relation "is the name of" does not demonstrate the required properties of an equivalence relation.

Sidman, Rauzin, Lazar, Cunningham, Tailby and Carrigan (1982)

Sidman, Rauzin, Lazar, Cunningham, Tailby and Carrigan (1982) again noted the difference between a conditional relation and "matching-to-sample". They suggested that identity matching is one case where the two relations may be confused. For, instance, given a vertical line sample stimulus the correct comparison may be another vertical line. While this relation may appear to be one of sameness, or identity, it may in fact be a conditional discrimination in just the

same way as if the stimuli were a vertical line sample and a colour comparison. In the same way, relations established between two physically different stimuli may be related only by conditionality or they may indeed be related by equivalence or "matching-to-sample". If the stimuli are related by equivalence they may have become mutually substitutable, as in the relation between the numeral 2 and the printed word TWO or the word RED and the hue red. Sidman et al suggested that it is tempting to assume that conditional discriminations establish equivalence as this would allow the use of conditional discriminations as a model for studying aspects of language, even with non-humans.

In both this paper and Sidman and Tailby (1982) the authors defined certain properties of the relationship between stimuli which should be demonstrable if the relations are truly equivalence relations - reflexivity, symmetry and transitivity. Sidman and Tailby demonstrated equivalence relations according to these criteria with normally able young children. Emergent equivalence relations were demonstrated with adolescents with learning disabilities (Sidman 1971, Sidman and Cresson 1973). Sidman et al noted however that they had been unable to demonstrate emergent equivalence relations rather than conditional discriminations with monkeys. The aim of Sidman et al was to describe some of the attempts to demonstrate symmetry with both rhesus monkeys and baboons.

In all experiments unless otherwise stated the subjects were adult rhesus monkeys.

Experiment 1 - The subjects were first given experience with line-line and hue-hue discriminations. This provided experience of all stimuli as both samples and comparisons. The subjects were then taught a line-hue conditional discrimination so that with vertical samples they learned to select green and with horizontal samples to select red. Finally the subjects were given probes for the emergent hue-line symmetrical relations.

The subjects' performances on baseline trials were above 90% correct but performances on test probes were around chance level. This suggested that training had established only a conditional relation between the sample and comparison stimuli. There was no evidence that the stimuli had become equivalent.

Experiment 2 - This experiment investigated the possibility that the incorrect comparison as well as the correct comparison may be important in specifying the conditional relations. In this case, the conditional relation in experiment 1 might

have been "If the sample is vertical and the comparisons are green *and red*, then green is the correct comparison" (Sidman et al 1982, p.30, italics in original). In this experiment, on the symmetry probes, the incorrect comparison remained the same and the sample and correct comparison were substituted for each other.

Even with the incorrect comparison held constant the subjects performance on the symmetry probes was around chance level. There was no improvement of performance from experiment 1. The conditional relations did not cause the stimuli to become equivalent.

Experiment 3 - A series of experiments had demonstrated that the conditional discrimination procedure does establish equivalence relations with a number of groups of subjects, e.g. retarded adolescents (Sidman 1971, Sidman and Cresson 1973, Sidman, Cresson and Willson-Morris 1973, Spradlin Cotter and Baxley 1973), normally able adults (Lazar 1977) and normally able children (Lazar and Kotlarchyk, note 2, Sidman & Tailby 1982). But these human subjects did not learn the same conditional relations as the monkeys and different methods of testing for equivalence were used. The aim of experiment 3 was to repeat experiment 1 with normally able children to see if the procedures would produce equivalence relations with this subject group.

The subjects in this experiment were six normally able children aged between 4:8 and 5:9 (years:months). Four of the six children demonstrated symmetry of the trained conditional relations. This suggested that the monkeys failure to demonstrate symmetry using this procedure was not due to some procedural artifact. However two of the children did not demonstrate symmetry, suggesting that the stimuli had not become equivalent. The training and test procedures were re-appraised and, following changes to the baseline conditional discriminations these two subjects did give positive results on the symmetry test. As the new baselines may have been responsible for this emergence of symmetrical responding, the new baselines were then tested with the monkeys in experiment 4.

Experiment 4 - Successful responding on the hue-line symmetry probes required a successive discrimination between the coloured samples and a simultaneous discrimination between the line comparisons. The results from experiment 3 suggested that the hue-hue and line-line training given may not have been sufficient to establish this repertoire. The hue-hue training did involve a successive discrimination between the hue samples but, when the subject was required to

select the correct hue comparison, both the red and green hues were available simultaneously as comparisons and so may not have required the subject to learn the successive discrimination as intended. Similarly the line-line training may not have established the simultaneous discrimination required for probe trials. For this reason, the line-line and hue-hue training trials were replaced with form-line and hue-form training to try and establish the respective simultaneous and successive discriminations (see Table 2:1).

Once again the subjects did not show appropriate symmetrical responding. The subjects tended to select the vertical comparison in the presence of both colour samples. New baseline discriminations had been introduced to ensure the subjects had experience with appropriate forms of conditional discriminations but these did not produce equivalence relations between the stimuli. Sidman et al also noted that the data from this experiment also revealed the absence of transitivity. The subjects were taught hue-form and form-line discriminations and these were sufficient to establish correct responding on the hue-line probes by transitivity, even without training the line-hue discrimination and testing for symmetry.

Experiment 5 - Different results were obtained using monkeys and children as subjects. It was not clear if this was due to species-related or experiential factors or both. Experiment 5 therefore repeated the procedure used in experiments 1 and 3 with 2 baboons as subjects.

Performance on probe-trials was around chance level for both baboons. This data strongly resembled the data obtained with the monkey subjects on this procedure. There was no evidence that training had established anything more than conditional relations between the stimuli. There was no evidence that the stimuli had become equivalent.

Sidman et al described a series of experiments which showed that procedures establishing conditional discriminations also established symmetrical responding with young children but not with monkeys or baboons. Experiment 4 also showed the absence of transitivity in the monkeys conditional discriminations. These results replicated and extended similar findings with pigeons (Hogan and Zentall 1977, Holmes 1979, Rodewald 1974). These procedures therefore, seemed to establish only conditional discriminations with monkey and baboon subjects but produced something more with the human subjects. For the children in this experiment the stimuli had become equivalent and it was appropriate to refer to

their performance as "matching-to-sample". Sidman et al pointed out that, as a result of the findings, it would not be appropriate to assume that training a conditional relation established anything more than a conditional relation, particularly with non-human subjects.

One reason for interest in conditional discrimination paradigms was that they might provide a model for studying language. "where the conditional discrimination procedure does generate true matching to sample, the formation of stimulus classes defines a semantic correspondence between each sample and its comparison stimulus" (Sidman et al, 1982, p.43). The failure of non-humans to generate matching to sample limited the use of this paradigm as a model. However, as Sidman et al pointed out, the analysis of non-human performance on these tasks was still potentially useful. Analysis of the reasons why non-humans failed to show equivalence performances might suggest the necessary components for this matching to sample performance. Techniques for establishing these components in the non-human behavioural repertoires might also have implications for humans with learning disabilities.

Summary from early experiments

These papers between 1971 and 1982 documented the investigation of the phenomenon that training some conditional relations could, with a number of groups of human subjects, lead to the emergence of new, untaught, conditional relations. Somehow these conditional relations had established more than conditionality, the stimuli involved had become equivalent - mutually substitutable.

These experiments investigating this finding showed a gradual change in how this was regarded and interpreted. The basic finding described by Sidman (1971) was considered in terms of reading and the pre-requisite skills for this behaviour, such as oral reading, auditory-receptive reading and reading comprehension. Much of Sidman (1971) and the other earlier papers (e.g. Sidman and Cresson 1973) was concerned with cross-modal transfer, for instance, how the visual behaviour of reading comprehension might have developed from previously learned auditory-visual equivalences. This approach seemed a logical one as it appeared to mimic developmental findings that visual behaviours such as reading emerged after auditory-visual behaviours such as selecting objects in response to spoken words. This shift, from auditory-visual behaviours to purely visual behaviours seemed to indicate a major developmental shift. Much of this early work then was aimed at

discovering what mediated this transfer as it might have important implications for assisting children who seemed unable to make this shift and were not developing behaviours such as reading.

One of the main aims then was to find what factors mediated this transfer and the data from these experiments were compared to results from "mediated-transfer" paired associate techniques such as those reported by Jenkins (1963) and Jenkins and Palermo (1964). However this emphasis gradually shifted to a more specific examination of the conditional discrimination paradigm itself. There was also a shift of emphasis from consideration of emergent behaviour in terms of reading and its pre-requisites (Sidman 1971; Sidman and Cresson 1973) and concepts (Spradlin, Cotter and Baxley 1973; Dixon and Spradlin 1976; Spradlin and Dixon 1976). Instead, the emphasis moved towards defining the precise conditions under which equivalence or "transfer" occurred in this conditional discrimination paradigm. VanBiervliet (1977) also described one experiment using this paradigm as a teaching tool.

This progression resulted in two very important papers: Sidman and Tailby (1982) and Sidman et al (1982). Sidman and Tailby's paper was important as it pointed out that it is not possible to assume that establishing a conditional discrimination will result in matching-to-sample or equivalence. Instead Sidman and Tailby described a series of properties which defined equivalence and proposed behavioural tests to determine if a conditional relation met these criteria for stimulus equivalence. This provided a basis for comparing across studies and ensuring a consistent evaluation of the phenomenon of emergent behaviour.

The stimulus equivalence paradigm in connection with reading had also developed considerably from Sidman (1971) to Sidman and Tailby (1982).

"Formal resemblances between conditional discriminations and reading do not prove one relevant to the other. The establishment of stimulus classes does prove the relevance. . . . The equivalence paradigm provides exactly the test that is needed to determine whether or not a particular conditional relation involves semantic relations" (Sidman and Tailby, 1982, p.20)

Equivalence was still seen as relevant to reading but in a different way . The stimulus equivalence paradigm was useful as a explanatory tool for behaviours such as reading rather than being a mediating factor in cross-modal transfer.

Sidman and Tailby (1982) also indicated how the stimulus equivalence paradigm had moved away from the idea of mediated-transfer. "The successful use of conditional discriminations to generate equivalences raises considerable doubt about the necessity for postulating the existence of mediating responses". (Sidman and Tailby, 1982, p.21). In the conditional discrimination paradigm the only overt response is pointing or touching, which is the same for all stimuli, therefore the transfer was not mediated by differential response topography. Sidman and Tailby had already stated that they believed there was evidence to show that stimulus equivalence was not mediated by naming. Therefore they concluded that "the very logic of the conditional-discrimination procedure suggests that no other kind of mediating response need be postulated" (Sidman and Tailby, 1982,p.22)

Sidman et al's (1982) experiments emphasised that it could not be assumed that a conditional discrimination procedure would necessarily produce equivalence. The procedure which produced symmetry with normally able children consistently failed to establish symmetry with non-human subjects. This emphasised the usefulness of Sidman and Tailby's (1982) definition of the necessary properties of equivalence relations. Testing for the properties of the equivalence relation provided a consistent definition of equivalence and also, if equivalence was not demonstrated, specified exactly where the relationship was breaking down. Sidman et al noted that examining why the conditional discrimination procedure establishes equivalence in humans but not in non-humans might be productive. Understanding why the procedure fails with non-humans might help show how it establishes equivalence with human subjects.

These experiments through the 1970's to 1982 showed efforts to describe and quantify the emergence of untaught relations. By 1982 this had produced a definition of the necessary properties of equivalence and specified tests for these properties, giving some indication of the conditions under which equivalence would or would not develop, and with which groups of subjects. The stimulus equivalence paradigm provided a way of examining naturally occurring emergent behaviour in a controlled laboratory setting. It was also becoming clear that equivalence had the potential to be a very powerful tool both for explaining and establishing new behavioural repertoires.

Fig. 2:1 Training and Test Relations, Sidman (1971)

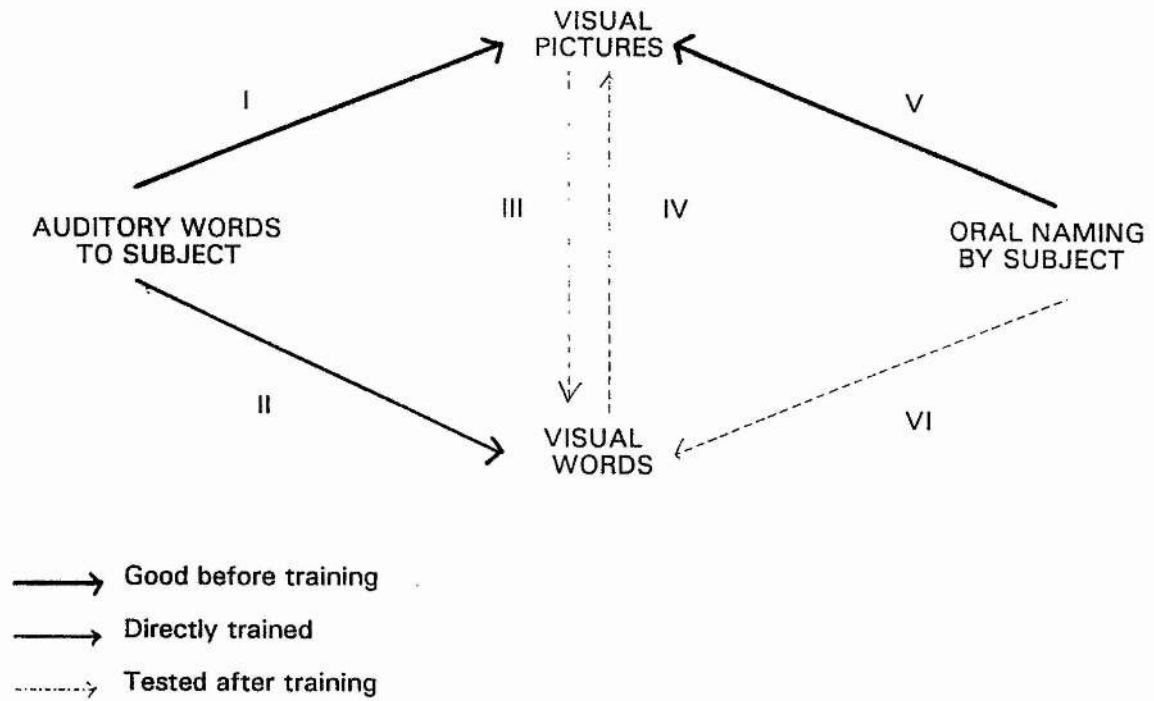


Fig. 2:2 Stimuli, Responses and Mediated-Transfer Paradigms.
From Sidman, Cresson & Willson-Morris (1974)

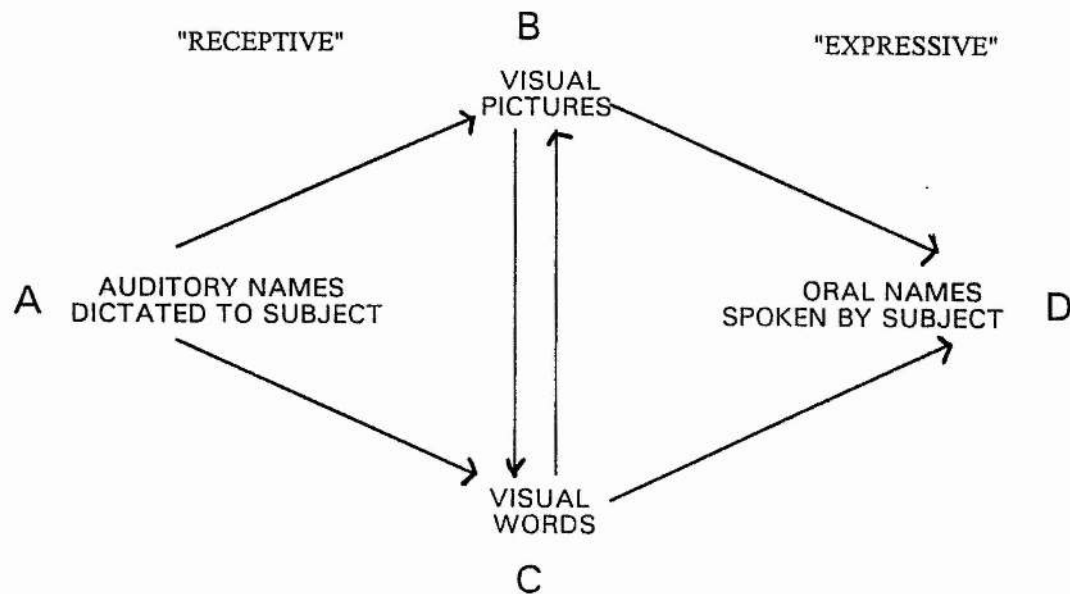


Fig. 2:3. Response Equivalence Paradigm, Jenkins & Palermo (1964)

Stimulus		Response
A	controls	B
A	controls	C
B	will tend to control	C
C	will tend to control	B

Fig. 2:4 Training and Test Relations, Van Biervliet, 1977

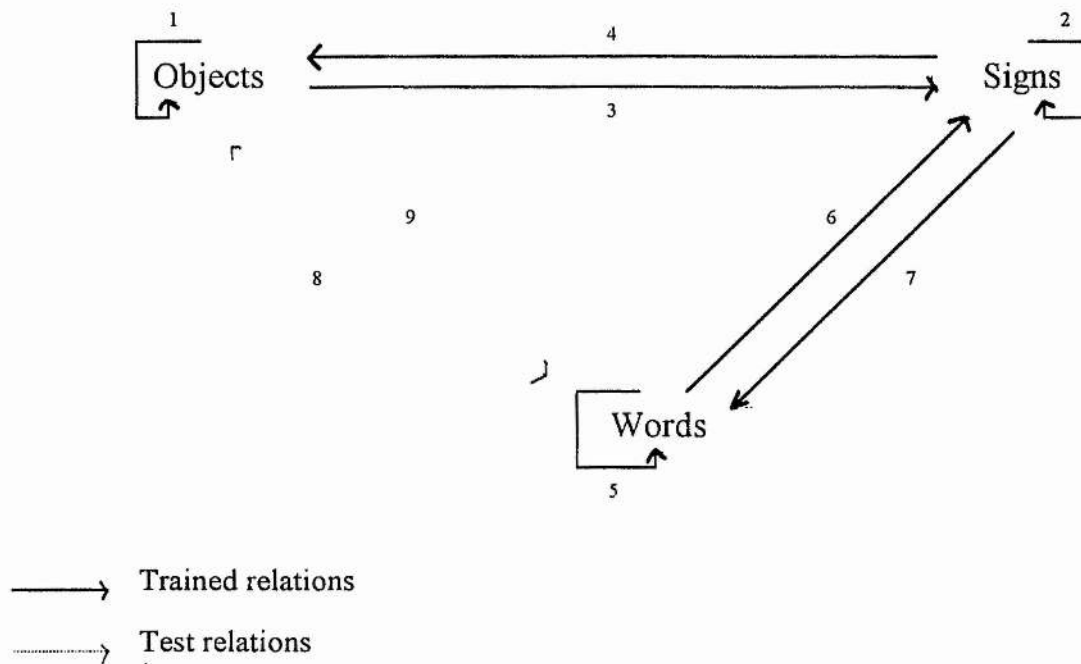


Fig. 2:5. Training and Test Relations, Adapted from Sidman and Tailby (1982)

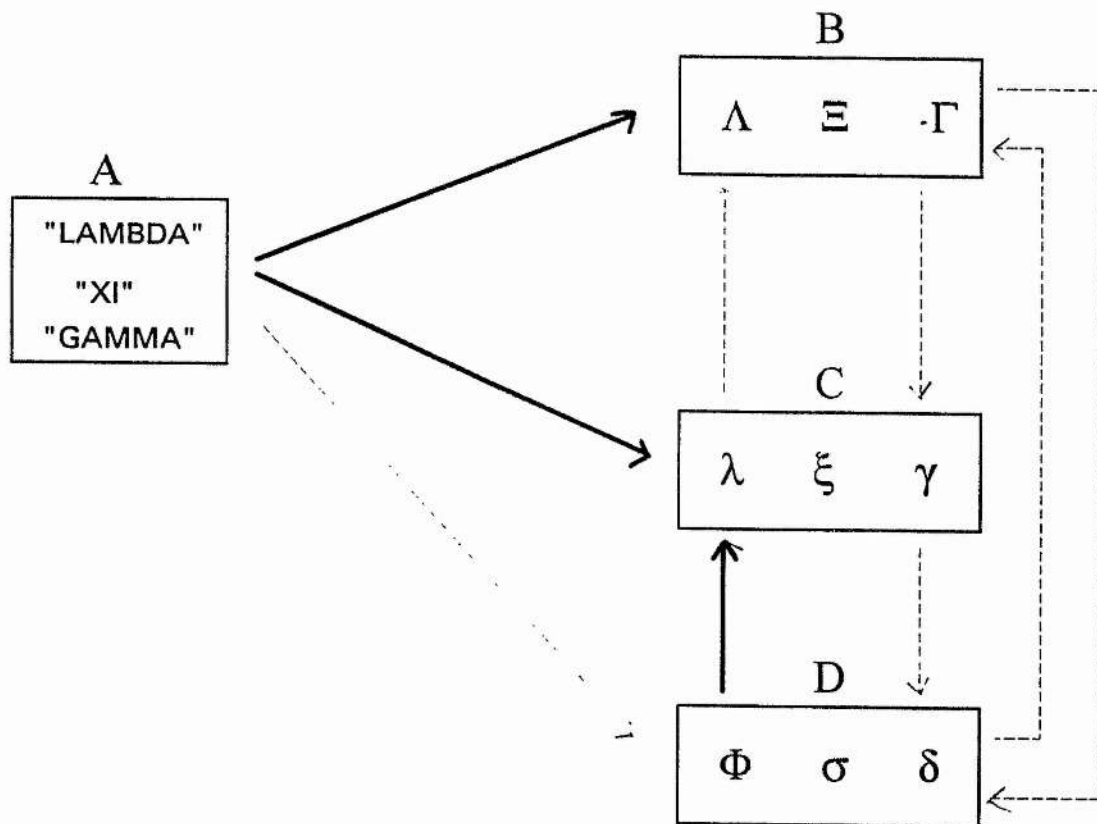


Table 2:1. Training and test relations, Sidman et al (1982)

	Sample	Comparisons	
		Correct	Incorrect
1A	x	vertical	horizontal
1B	+	horizontal	vertical
2A	green	x	+
2B	red	+	x
* 3A	vertical	green	red
* 3B	horizontal	red	green
4A	green	vertical	horizontal
4B	red	horizontal	vertical

Trials 1A & B - form-line, trials 2A & B hue-form.

Trials 3A & B - trained conditional discrimination

Trials 4A & B - symmetry probes

CHAPTER 3

Examples of Cognition in Non-Humans

The Historical Background

There has long been interest in the cognitive abilities of non-humans, and particularly in how their abilities might relate to the cognitive abilities of human subjects on similar tasks. "Since the turn of the century the experimental investigation of problem solving has been a fruitful source of information on how animals and men think and learn" (Scheerer, 1963, p.26)

Romanes (1888)

Even before the experimental investigation of non-human cognitive abilities attempts were made to describe animal intelligence. Romanes (1888) considered the reliability of various accounts of animal behaviour, from both scientific sources and from popular literature. Romanes suggested that it was not possible to accept many of these accounts uncritically, as many were from little known researchers, or even anonymous. Yet Romanes also noted that accepting reports only from well established researchers might neglect some valid evidence from lesser known observers. Romanes therefore suggested three principles to be applied when considering reports of animal intelligence. Firstly, not to accept any alleged fact without the authority of some respected name. Secondly, if an alleged fact seemed important but the person reporting it was not well known, it was important to consider if there was much opportunity for mal-observation. In other words, for the fact to be accepted the behaviour of the animal had to be clear and unmistakable. Thirdly, Romanes tabulated all observations by unknown observers in order to check if reports by one observer were corroborated by other independent observers. Romanes felt that this corroborated evidence was as trustworthy as statements based on the authority of a single, though well respected observer.

At this time, consideration of non-human cognitive abilities tended to be based on reports of animal behaviour rather than experimentation. However, Romanes documents principles for testing the quality of these reports and bases the limits of animal intelligence only on reports with considerable authority or on behaviours which had been clearly demonstrated on several occasions.

Romanes pointed out that when considering the phenomenon of "mind" in relation to non-humans only *objective* analysis is possible. Subjective analysis is possible of

only a person's *own* mind as only that person has immediate cognisance of the processes operating. For all other minds, both human and non-human, such immediate cognisance is not possible and all knowledge of the processes operating can only be inferred through the medium of what Romanes referred to as ambassadors - the activities of the organism. From these observable activities the intention is to infer the underlying mental operations. However, as Romanes then pointed out, it can be hard to say exactly which abilities really show evidence of intelligence or "mind" - even quite complex adaptive behaviours can be performed without true purpose or consciousness. Romanes defines behaviour giving evidence of intelligence as adaptive behaviour where the mechanism of the nervous system, or reflex behaviour, is not sufficient to explain the behaviour observed.

"Does the organism learn to make new adjustments, or to modify old ones, in accordance with the results of its own individual experience?. If it does so, the fact cannot be due merely to reflex action . . . for it is impossible that heredity can have provided in advance for innovations upon, or alteration of, its machinery during the lifetime of a particular individual" (p.46)

Romanes goes on to point out that this definition of an organism showing itself able to learn by its own experiences requires the same evidence in animals as in humans. If this learning by experience is sufficient proof for evidence of mind/intelligence in humans then it must also be accepted as sufficient proof for evidence of mind/intelligence in non-humans.

Romanes went on to give some examples of observations which seemed to show that non-humans, and particularly monkeys and apes, displayed behaviour consistent with this definition of intelligence. For instance, a dog will eat when hungry, a reflex action. But the dog can also learn not to eat, even when hungry, until it has received a particular verbal signal. Romanes suggests that this shows the same evidence that the dog's actions are prompted by mind as would be evidence with a human subject. Romanes also described several instances of apes learning to use sticks, and other objects, to obtain items which were out of reach, and suggested that these also fitted with a definition of learning by experience.

Thorndike (1898)

Thorndike (1898) however argued that mere descriptions of behaviour are an insufficient basis with which to consider intelligence in non-humans. Thorndike states that "The main purpose of the study of the animal mind is to learn the development of mental life down through the phylum, to trace in particular the origin of human faculty" (p.63). However, Thorndike felt that this study of the

animal mind was poorly served by a descriptive account of animal behaviour. He argued that descriptions of animal behaviour "have all been about animal *intelligence*, never about animal *stupidity*" (p.64, italics in original) and were thus prejudiced to find evidence of intelligence in animals. Thorndike believed that this prejudice was worsened by the "facts" about animal intelligence being drawn from observations and anecdotes. He suggested that there was a tendency to report cases of exceptional animal behaviour while neglecting normal animal behaviour; "In short, the anecdotes give really the *abnormal* or super-normal psychology of animals" (p.65, italics in original). Even applying Romanes criteria of the report having the authority of a respected name and ruling out the possibility of mal-observation is of little use if there is a prejudice towards reporting atypical behaviour. Possibly only Romanes third criterion of accepting reports of the same behaviour from independent researchers goes some way to addressing this problem.

Thorndike stated that the solution to the problem was to investigate animal intelligence through experimentation rather than observation. He felt that this procedure had notable advantages to just observation; the conditions could be repeated on several occasions to ensure that the animals behaviour on the first occasion was how it would typically behave, and a number of animals could be given the same test to ensure that the behaviour shown was typical of the species. Thorndike applied these principles by constructing experiments with a number of animals (cats, dogs, chicks) where the animal was put in a box or pen from which it could free itself, and gain access to food or company, by some simple device such as pressing a lever or pulling at a loop of cord. The animals behaviour in this situation was then observed - did the animal show any evidence of insight into the problem or could its eventual release have been explained purely in terms of trial and error. Thorndike suggested that with this sort of approach

"The animals are put in situations which call into activity their mental functions and permit them to be carefully observed And this actual vision of animals in the act of using their minds is far more fruitful than any amount of histories of what animals have done without the history of how they did it" (p.67).

Even more importantly, the animals behaviour was free of any influence by the observer. The only personal factors, according to Thorndike, were in the observation and interpretation and as the precise conditions of the experiment were recorded the experiment could be repeated independently by other researchers.

Lloyd Morgan (1909)

This experimental approach advocated by Thorndike proved to be useful for the analysis of both human and non-human cognition. When it came to interpretation of the experimental data, as Scheerer (1963) has pointed out, one of the most influential views was that of Lloyd Morgan (1909) "In no case may we interpret an action as the outcome of the exercise of a higher psychical faculty, if it can be interpreted as the outcome of the exercise of one which stands lower in the psychological scale" (p.56). Or as Scheerer puts it "Always explain behaviour in terms of the simplest possible mechanism" (Scheerer, 1963, p.26). As an example of this, Lloyd Morgan describes his dog carrying a stick which was much heavier at one end than the other. Initially the dog held the stick in the middle but this proved to be unbalanced and awkward as one end of the stick was much heavier than the other. However, after some experience with the stick the dog began to hold the stick much closer to the heavier end which made the stick more balanced and easier to carry. As Lloyd Morgan pointed out, by this behaviour the dog had found the centre of gravity of the stick - a practical solution to a mechanical problem. However, when considering the dogs behaviour there is no need to assume that the dog had any perception of the spatial relationship of the stick and the centre of gravity. Instead, according to Lloyd Morgan's canon of explaining behaviour in the simplest possible terms, the dog's action can be explained by "sense-experience". By trial and error the dog found the most comfortable way of carrying the stick, and while this was an intelligent adaptation to the circumstances there is no need to assume higher processes of insight or perception of the relations between the form of the stick and its centre of gravity. This particular behaviour can be explained in terms of the lower process of sense-experience rather than the higher processes of perception or insight.

This account of the dog's behaviour is similar to the conclusion reached by Thorndike as a result of his experiments "Whatever behaviour is rewarded is "stamped in " and whatever behaviour is not rewarded is "stamped out"" (Scheerer, 1963, p.27). This explanation does not require assumption of intelligence or insight, merely that the animal is capable of learning new habits.

Kohler (1927)

There were some objections to Thorndike's assessment of the animals behaviour in his experiments, particularly by the Gestalt psychologists. Thorndike's animals gained release from their enclosure by a process of trial and error, but Kohler argued that this was because the experimental setting did not permit the animal to behave intelligently. Kohler contrasted the behaviour of the animals in Thorndike's

experiments with the behaviour of a chimpanzee in a related setting. In Kohler's experiment a hungry chimpanzee was confined in a cage. A banana was outside the cage just beyond the chimpanzee's reach, and there was a stick lying on the floor of the cage. The chimp made a few attempts to reach the banana through the bars of the cage but was unsuccessful. In one action, the chimp then picked up the stick and used it to retrieve the banana. This swift solution was in contrast to the trial and error behaviour shown by the animals in Thorndike's experiment. Kohler (1927) asserted that the chimp was able to behave intelligently in this situation because the solution was part of the overall situation (one of the defining features of Gestalt psychology) and so the chimp was able to perceive the solution to the distance problem. Kohler argued that Thorndike's test setting did not permit this sort of intelligent behaviour by the animal as the release mechanism in the cage was hidden and therefore not part of the perceived situation, and in any case its functioning was too complicated for an animal to comprehend.

What is notable though is that Kohler had no objection to Thorndike's assertion that experimental investigation of animal behaviour is necessary. Rather, this debate seems to prove the point Thorndike was making - by controlling the setting for the behaviour it is possible to identify the causes of the behaviour. Different experimental settings between Thorndike's and Kohler's experiments produced different behaviours and further manipulation of these variables was one way of establishing the limits of non-human cognitive abilities.

Stimulus Equivalence

There have been many attempts to demonstrate "human" cognitive abilities with non-human subjects. Numerous attempts have been made to demonstrate stimulus equivalence with non-humans, and to clarify under what, if any, conditions it is possible to demonstrate equivalence with these subjects. Sidman and Tailby (1982) defined the necessary properties of an equivalence relation; reflexivity, symmetry and transitivity. Sidman et al (1982) attempted to demonstrate one of these properties, symmetry, with both rhesus monkey and baboon subjects. The evidence from Sidman et al (1982) suggested that even if non-humans can demonstrate equivalence relations, they do not demonstrate them with the same ease as human subjects. Certainly the procedures used by Sidman et al (1982) which produced equivalence with normally able children gave no evidence of establishing stimulus equivalence with the non-human subjects (see chapter 2 for fuller discussion).

Since Sidman et al (1982) there have been several attempts to demonstrate stimulus equivalence as defined by Sidman and Tailby (1982) with non-humans. Generally these studies have had little success, both with pigeon and primate subjects.

D'Amato, Salmon, Loukas and Tomie (1985)

This experiment by D'Amato et al followed on from the work by Sidman et al (1982) testing for symmetry and transitivity with non-human subjects. Sidman et al (1982) had found no evidence for symmetry or transitivity with non-human primate subjects. D'Amato et al (1985) tried to verify Sidman et al's results with monkey and pigeon subjects. D'Amato et al found no evidence for backward associations (symmetry) following conditional discrimination training with monkeys, "If monkeys form backward associations . . . those associations tend to be weak and easily swamped by competing sources of control" (p.41). However, following training of further conditional relations, the same monkey subjects showed convincing evidence of associative transitivity. D'Amato et al concluded that "associative transitivity appears from the present results to be a robust phenomenon in monkeys" (p.44). D'Amato et al then replicated this transitivity experiment as closely as possible with pigeon subjects. However there was no evidence of associative transitivity with the pigeons.

D'Amato et al therefore found evidence for one of the properties of stimulus equivalence, transitivity, with one set of non-human subjects, monkeys. However there was no evidence for symmetry with the same group of subjects and no evidence for transitivity with a different set of non-human subjects, pigeons.

Dugdale and Lowe (1990)

Dugdale and Lowe (1990) also reported an experiment evaluating symmetry with non-human subjects. Dugdale and Lowe's study is made particularly interesting by the subjects they used. Dugdale and Lowe had hypothesised that linguistic ability, and in particular naming, are necessary in order for a subject to demonstrate emergent equivalence relations, and this might explain why it had proved so difficult to demonstrate equivalence with non-human subjects. Dugdale and Lowe investigated the possibility of demonstrating stimulus equivalence with non-human subjects who had received language training. These subjects were three chimpanzees who had all participated in an ape-language training programme and who could communicate by touching lexigrams on a key board, each of which was

associated with an object, action or location. (Rumbaugh 1977, Savage-Rumbaugh 1984).

Dugdale and Lowe attempted to teach the chimps an AB matching relation and then to test for the emergence of the BA symmetry relation. One subject failed to learn the AB matching task. The other two subjects learned the AB task but showed no evidence of BA symmetry, despite several procedural modifications designed to maximise the chimpanzees chances of success. These modifications included adding identity matching trials to the AB baseline so that the subjects had experience with all the stimuli as both samples and comparisons, and also providing reinforcement for correct responses on both baseline and test trials. Despite these modifications, neither subject showed any evidence of emergent BA symmetrical relations.

Dugdale and Lowe argued that, given the failure to display one of the properties of equivalence with these subjects, it was likely to be very difficult to display stimulus equivalence with other non-human subjects. Chimpanzees seemed more likely candidates to demonstrate equivalence than more distant non-human relatives such as pigeons and these particular chimpanzee subjects had an unprecedented history of complex training and testing. However, it must be noted that Dugdale and Lowe's procedure failed to establish even AB matching with one subject, and in their review of stimulus equivalence with non-verbal organisms Dube, McIlvane, Callahan and Stoddard (1993) suggested that limitations in experimental procedures may have contributed to failures to demonstrate aspects of equivalence relations, at least with rats and pigeons.

Both D'Amato et al (1985) and Dugdale and Lowe (1990) showed failures to demonstrate some of the crucial components of equivalence with non-human subjects and other studies have found similar results (e.g. Kendall 1983, Lipkens, Kop and Mathijs 1988, Sidman et al 1982)

Two studies did claim to have demonstrated stimulus equivalence with non-human subjects: McIntire, Cleary and Thompson (1987) and Vaughan (1988). However, several researchers have argued that it is inappropriate to regard the performances of the non-human subjects on these tasks as demonstrating stimulus equivalence.

McIntire, Cleary and Thompson (1987)

McIntire Cleary and Thompson suggested that possibly two stimuli could be related because the same characteristic response is generated in the presence of each. McIntire et al suggested that humans readily engage in behaviours such as this, e.g. emitting verbal names. By training monkeys to emit a characteristic responses for stimuli within a set, McIntire et al hoped to demonstrate equivalence relations with non-humans. The monkeys were trained to respond to two sets of three stimuli. Stimuli 1, 3 and 5 were set "odd" and stimuli 2, 4 and 6 were set "even". A topographically different response, R, was trained to each set; R(odd) and R(even). An "odd" response required the subject to press the stimulus key and hold that press continuously for 3.5 seconds. An "even response requires the subject to press and release the stimulus key eight times successively. In a training trial the animals were required to emit the appropriate topographical response to the sample and then to the correct comparison, any other combination of stimulus selections or response topographies was not reinforced.

McIntire et al trained six relations (notation from Hayes 1989):

- A. 1-R(odd) - 3-R(odd)
- B. 3-R(odd) - 5-R(odd)
- C. 1-R(odd) - 1-R(odd)
- D. 2-R(even) - 4-R(even)
- E. 4-R(even) - 6-R(even)
- F. 2-R(even) - 2-R(even)

For example, in relation A, the subject had to emit response R(odd) to sample stimulus 1. This produced two comparison stimuli, one "odd" and one "even". The subject then had to correctly select comparison stimulus 3, and emit the appropriate R(odd) response to that stimulus.

McIntire et al used two monkey subjects. Following training with the six relations detailed above, McIntire et al claimed that both subjects demonstrated the emergent properties of reflexivity, symmetry and transitivity and both passed the equivalence test requiring a combination of symmetry and transitivity. McIntire et al stated that these performances were robust and stable in the absence of reinforcement.

However this procedure has been criticised on the grounds that it directly trains all the components of the tested relations, and as no relations were derived, the subjects did not show true stimulus equivalence. Hayes (1989) has argued that if relations A, B, D and E alone had been trained, McIntire et al's findings would

have been evidence for equivalence in non-humans, but that by training relations C and F, McIntire et al directly trained all parts of the relations subsequently tested. For example, in relation C, 1-R(odd) - 1-R(odd), the animal sees stimulus 1 as the sample and emits response R(odd). Given stimuli 1 and 2 as comparisons the animal is reinforced for selecting stimulus 1 using response R(odd) rather than stimulus 2-R(even). The animal is thus reinforced for selecting stimulus 1 immediately after engaging in the R(odd) response. A similar situation exists following training of relation F, 2-R(even) - 2-R(even), where the animal is reinforced for selecting stimulus 2 immediately after engaging in the R(even) response. During the test phase of the experiment the test for combined symmetry and transitivity requires the subject to select stimulus 1 as the correct comparison when stimulus 5-R(odd) is the sample, and select stimulus 2 as the correct comparison when stimulus 6-R(even) is the sample. However, as Hayes (1989) notes, the animal has been reinforced for selecting stimulus 1 after engaging in an R(odd) response and selecting stimulus 2 after engaging in an R(even) response. Therefore the animal can respond correctly on the tests for equivalence without actually deriving any of the relevant relations. Instead, the animals performance on this task can be explained in terms of reinforced chaining. Hayes (1989) details how the components of the test relations of symmetry, reflexivity and transitivity have been directly reinforced in two or more of the training relations. Hayes states that

"Because a name has been directly trained to both precede and follow all stimuli of the same set, moving from one to another in any possible test trial can occur as a series of simple discriminations. Nothing is derived about this performance" (p.387)

Because the test performances were not derived, the performances by McIntire et al's non-human subjects do not satisfy the conditions for stimulus equivalence according to Sidman and Tailby (1982).

Another study which claimed to have demonstrated stimulus equivalence with non-humans was a study by Vaughan (1988)

Vaughan (1988)

Vaughan examined stimulus equivalence in a different way to Sidman and Tailby (1982) and other researchers who trained conditional relations and tested for the emergence of reflexivity, symmetry and transitivity. Vaughan suggested that there was another way of viewing an equivalence relation; a partition. Vaughan defined a partition as consisting of "a set of disjoint subsets of the set, with the union of those

sets equalling the original set . . . Any equivalence relation defined on a set induces a partition of that set, and any partition implies an equivalence relation" (p.36). For instance, Vaughan suggested that a similarity relation could be one example of an equivalence relation. In this case a similarity relation would partition a set of stimuli so that all the stimuli within a subset are similar to each other, while no stimuli in different subsets are similar to each other.

Vaughan investigated equivalence in the form of a partition using pigeons as subjects. Forty different slides of trees were assigned randomly to two sets of twenty slides. One set was arbitrarily defined as S+ and the other set as S- and the two sets were presented mixed together. When performance was stable so that responses were high for S+ slides and low for S- slides, the functions of the stimuli were reversed so that the S+ slides became S- and the S- slides became S+. When the reversal discrimination was established and performance was stable, the contingencies were reversed once more. This discrimination training continued until eventually exposure to the first few slides in a set following reversal of the contingencies led to appropriate performances with the remaining slides in the set.

This study by Vaughan (1988) showed that functional stimulus classes can be established with non-human subjects. What has been disputed however, is whether Vaughan's study demonstrates anything more than the establishment of functional stimulus classes; in particular, is this definition of stimulus equivalence really relevant to the definition of stimulus equivalence by Sidman and Tailby (1982). Hayes (1989) has pointed out that this procedure does not allow for tests for the necessary properties of reflexivity symmetry and transitivity. This is part of the fundamental issue that a critical aspect of equivalence is the emergence of performances not traced to a history of direct reinforcement. Hayes (1989) argued that the contingency class performances had been directly trained through the repeated reversal procedure. Hayes therefore believes that Vaughan's procedure had established only functional stimulus classes rather than equivalence classes; "Equivalence relations are defined by the derived relations among members. Functional classes are defined by common response functions - nothing need be derived" (Hayes 1989 p.390).

Dube, McIlvane, Callahan and Stoddard (1993) agreed with Hayes (1989) asserting that Vaughan's study did not demonstrate stimulus equivalence (However see chapter 5 for fuller discussion of Vaughan 1988). Nevertheless, Dube et al did feel that Vaughan's study was significant as it provided a convincing demonstration of

functional stimulus class formation with non-human subjects and such class formation had previously only been demonstrated with humans. Dube et al (1993) pointed out that several studies with human subjects had shown that contingency class membership could be independent of equivalence class membership and so Vaughan's (1988) demonstration of contingency classes with non-humans could not be taken as a demonstration of equivalence classes with non-humans. However Dube et al also noted that contingency class membership was frequently accompanied by equivalence membership, suggesting that the two forms of class membership may have common behavioural prerequisites. Therefore, although Vaughan's findings can not be taken as a demonstration of equivalence class formation with non-humans, these findings are relevant to the question of whether it is possible to demonstrate stimulus equivalence with non-human subjects.

While neither McIntire, Cleary and Thompson (1987) nor Vaughan (1988) demonstrated true stimulus equivalence with non-human subjects, there was some evidence that aspects of stimulus equivalence such as symmetry or transitivity independently could be shown with non-humans (e.g. Boysen & Berntson, 1989; D'Amato et al, 1985, Savage-Rumbaugh, Rumbaugh, Smith & Lawson, 1980; Tomonaga, Matsuzawa, Fujita & Yamamoto, 1991) and Vaughan's (1988) study had demonstrated functional stimulus classes with non-humans. However it was also clear that non-humans did not demonstrate stimulus equivalence with the same ease shown by a variety of human subjects. Dube et al (1993) analysed two possible interpretations for this difference between humans and non-humans on this task. They suggested that the difference could be either qualitative or quantitative. They suggested that if the difference between human and non-human performances was *qualitative* this would suggest fundamentally different behavioural processes were operating. If this were true it may require an explanation in terms of associative neural systems and such networks would be most appropriately investigated by biobehavioural techniques such as neuroimaging. On the other hand, the difference might be *quantitative*, implying that the failure to demonstrate equivalence with non-humans may be due to procedural deficits such as inadequate preparation, inadequate testing, or both. Dube et al suggest that a quantitative interpretation implies a behaviour-analytic investigation is most appropriate.

"The standard equivalence procedures involves behavioral pre-requisites that humans, but not laboratory animals are likely to have acquired through pre-experimental experience. . . . The research path is to identify the behavioral pre-requisites for equivalence by developing an instructional program that will *teach*

equivalence by supplying those pre-requisites." (Dube et al, 1993, p.763, italics in original)

Quantitative approaches had already been used in this way to extend the range of behaviours shown by non-humans.

Dube et al's (1993) suggestion that it may be necessary to identify and teach behavioural pre-requisites in order to demonstrate equivalence with non-humans seems to be supported by a study carried out by Schusterman and Kastak (1993). This study appears to demonstrate stimulus equivalence quite convincingly with a sea-lion subject.

Schusterman and Kastak (1993)

Schusterman and Kastak (1993) noted that two studies not long prior to their own study had either failed to show symmetry with chimpanzee subjects (Dugdale and Lowe, 1990) or showed it only weakly (Tomonaga et al, 1991). Schusterman and Kastak suggested that one reason these studies failed to demonstrate symmetry was that the subjects did not have sufficient experience of the sample and comparison stimuli switching roles prior to equivalence testing with novel symmetrical relations. They also suggested that performances on test trials were subject to extinction procedures, which may have affected the results.

In their study Schusterman and Kastak used a simple-to-complex training protocol similar to that developed by Adams, Fields and Verhave (1993). In this simple to complex protocol, the test for each property of equivalence is given immediately after demonstration of its pre-requisites so, for example, accurate performance on AB is the only necessary pre-requisite for the BA symmetry test. Subsequent training of BC allows for testing of CB symmetry and AC transitivity. During simple-to-complex training, if the subject fails to reach criterion on any test, that step is explicitly taught before the subject progresses to the next step. Schusterman and Kastak believed that this extensive training provided a breadth of experience that was important in allowing the subject to develop equivalence relations, and that training of this sort was most likely to facilitate equivalence class formation.

The procedure used in this experiment had several features designed to maximise the chances of a non-human subject demonstrating equivalence relations. Thirty potential 3 member classes were trained. The first twelve were used to provide the sea-lion subject with experience of sample and comparison stimuli switching roles. The other 18 were to be used for the CA equivalence test. This sample size was

large enough to avoid using extinction procedures. Instead the first test trial was critical. To pass tests of symmetry, transitivity and equivalence, the subject had to respond correctly on the first test trial and also had to respond correctly on at least three out of the first four trials of a problem. Correct performance on trial one was used to show that the subject's performance was not dependent on reinforcement and the data from further trials was to show that this performance was relatively reliable (although reinforced). The subject's ability to infer reflexive relations was also checked with tests of generalised identity matching.

Schusterman and Kastak began testing with two sea-lion subjects. One subject failed to maintain the AB and BC relations and so data was reported from only one subject. This subject (Rio) was taught AB relations with all 30 potential classes. Twelve classes were then removed for testing and training. These classes were taken at random from the early, middle and late phases of training. These twelve classes were then tested for the emergence of BA symmetry, in two groups of six. If necessary, the BA relations were trained to criterion. The BC relations were then trained for all 30 potential classes. The same 12 classes were then removed and tested for CB symmetry and if necessary these relations were trained to criterion. These 12 classes were then tested for AC transitivity and again these relations were trained if necessary. Finally, these 12 classes were tested for CA equivalence and if necessary these relations were trained. Finally, the remaining 18 potential classes were tested for CA equivalence. This tested the subjects ability to combine transitivity and symmetry relations without having had previous symmetrical or transitive experience with them.

Following AB training, Rio's performance on the BA symmetry test with the 12 selected relations was not significantly better than chance, although she did show some improvement in the second test session relative to the first. Following BC training, performance on the CB symmetry test was shown to be significantly above chance level. On the AC transitivity test Rio passed 11 out of 12 transitivity problems, a performance which was shown to be significantly better than that expected by chance. Following these tests of AB, BC and AC relations, Rio's performance on the CA symmetry test was again significantly better than that expected by chance.

The important equivalence tests were those carried out with the remaining 18 potential classes. Rio reached criterion on 14 out of 18 test problems and made only two first-trial errors out of 18 presentations. These results were shown to be

significantly better than would have been expected by chance. This suggested that a non-human subject had in fact demonstrated stimulus equivalence following training on AB and BC conditional relations.

Kastak and Schusterman (1993) showed that California sea-lions could transfer identity match-to-sample performance to novel visual stimuli. This study by Schusterman and Kastak (1993) extended this finding to examine the ability of a sea-lion to form concepts of symmetry and transitivity following conditional discrimination training. On the initial BA symmetry test, Rio's performance failed to reach criterion, a finding which was consistent with previous studies of symmetry with non-humans. Sidman et al (1982) failed to demonstrate symmetry with non-humans, and they suggested that providing enough examples might facilitate the emergence of symmetry with non-humans. The data from Schusterman and Kastak seems to support this suggestion. Schusterman and Kastak state "We believe the critical factor in Rio's subsequent performance in passing tests of symmetry, transitivity and equivalence stems directly from her experiencing enough exemplars to group these interrelated concepts" (p.836). A study by Wright, Cook, Rivera, Sands and Delius (1988) had also shown multiple exemplar training to be helpful in acquisition of identity relations with pigeons. It also seems likely that the simple-to-complex training and test protocol was helpful in establishing equivalence performance by ensuring establishment of all the pre-requisite repertoires.

Hayes (1989) had suggested that it might be harder to show derived symmetry with a non-human than to show derived transitivity. Symmetry requires a bi-directional stimulus-stimulus relation while transitivity only requires a uni-directional stimulus-stimulus relation although transitivity does require the co-ordination of two terms by means of a third, linking term. In this case, it may be easier to demonstrate accurate non-human performance on a task based on transitivity rather than symmetry. For this reason, experiments on transitive inference tasks using non-human subjects may be helpful in considering the range of animal cognition. Like equivalence, transitive inference requires the derivation of new relations on the basis of previously learned relations, but given Hayes' (1989) comments it would be expected that it would be easier for non-humans to demonstrate derived relations on a task of this sort.

Transitive Inference

Transitive inference provided a useful way to explore the limits of non-human cognition. Higa and Staddon (1993) suggested that there has been a tendency by cognitive psychologists to treat humans as isolated subjects and ignore comparisons between human and non-human abilities. They believed that this isolated approach was mistaken and that transitive inference provided a way to "study" the conditions under which behaviour that in humans would be symbolic can be brought about, by non-verbal means in other animals" (p.265). McGonigle and Chalmers (1977) also believed that the transitive inference task could be of use to comparative psychology. They pointed out that monkeys seemed able to learn complex tasks but it was not clear if they could reason. One problem was that the reasoning tests used tended to be very verbally dependent and "it is difficult to devise tests which are both meaningful to non-verbal subjects yet satisfy the stringent requirements of a formal reasoning test" (p.694). These formal reasoning tests tended to be very like those used by Burt, and later Piaget:

1. Edith is fairer than Suzanne
2. Edith is darker than Lili
3. Who is the darkest, Edith, Suzanne or Lili ?. (Burt 1919)

However, a transitive inference test for non-humans became possible following work by Bryant and Trabasso (1971).

Bryant and Trabasso (1971)

Piaget had suggested that children could not form transitive inferences until the age of about seven, when they reached the stage of concrete operations. Bryant and Trabasso suggested that children much younger than seven could in fact make transitive inferences provided they understood, and could remember the items of information they were being asked to combine. Bryant and Trabasso therefore suggested a simpler way of presenting the relational information the children had to combine in order to make a transitive inference. They taught the children a series of four overlapping pair discriminations, $A > B$, $B > C$, $C > D$, $D > E$, and then presented the test pair BD . By combining $B > C$ and $C > D$ the children could then infer $B > D$, even though the relation between these two stimuli had never been directly stated. The series of five stimuli were necessary in order to control for absolute responding. If a three stimulus series had been used, $A > B$, $B > C$, the subject could respond correctly to A rather than C without making a transitive inference just because responses to A are always correct and responses to C are never correct. Using a five term series, $A > B > C > D > E$, prevents this as the stimuli in the critical

test pair BD are sometimes correct, $B > C$, $D > E$, and sometimes incorrect, $A > B$, $C > D$. Therefore the subject must make a transitive inference in order to derive $B > D$.

Bryant and Trabasso trained the overlapping series of pairs by presenting their normally able child subjects with different coloured wooden rods of differing lengths so that rod A was long and rod E was short. These were presented to the children in pairs embedded in a wooden base so that only the top inch of each rod was visible. This meant that the children had to use the different colours of the rods to make a choice between the different lengths. Using this procedure Bryant and Trabasso were able to demonstrate transitive inferences with children as young as four.

The children in this experiment were asked to make a choice on the basis of which rod they believed was longer, so Bryant and Trabasso's procedure still had a verbal component.

McGonigle and Chalmers (1977)

McGonigle and Chalmers (1977) adapted Bryant and Trabasso's procedure to a non-verbal procedure and tested for transitive inference with squirrel monkeys. In this case the stimuli were containers of equal sizes but different colours. When the stimuli were presented in pairs, one stimulus was designated "light" and was empty, and the other was "heavy" and filled with lead shot, so that, for example, the rewarded stimulus was heavier than the non-rewarded stimulus. As only the two weights of "light" and "heavy" were used, no specific weight could be identified with stimulus B, C or D, so that when given the BD test pair the subjects could only respond correctly on the basis of the relational information trained. To select a stimulus the monkey had to move the tin from its position to reveal a foodwell. A peanut was concealed in the foodwell under the stimulus designated correct. Using this procedure McGonigle and Chalmers were able to demonstrate transitive inference performances with seven out of eight monkey subjects. The eighth subject did not show sufficiently stable training performance to be tested. The transitive choice performances shown by McGonigle and Chalmer's squirrel monkeys was very similar to the performances shown by Bryant and Trabasso's normally able four year olds.

This basic finding of McGonigle and Chalmers (1977) with non-human subjects has been replicated on several occasions with different subjects. Gillan (1981)

repeated the procedure with a chimpanzee subject, Fersen, Wynne, Delius and Staddon (1991) used pigeons as did Higa and Staddon (1993), and Davis (1992) demonstrated transitive inference in rats. All these studies demonstrated, to a greater or lesser degree, good evidence of transitive inference performance with non-human subjects.

Several of these studies also manipulated the transitive inference training paradigm to investigate further the precise conditions under which subjects would demonstrate transitive inference. This involved removing the end points from the transitive inference series and seeing what effect this had on transitive inference responding.

Gillan (1981)

This manipulation was first carried out by Gillan (1981). Using a chimpanzee subject, Sadie, Gillan trained the overlapping pairs A-B+, B-C+, C-D+, D-E+ (on each trial the subject received reinforcement for selecting the stimulus indicated +, and did not receive reinforcement for selecting the stimulus indicated -) to establish the series $A < B < C < D < E$ (Experiment 1A). Following this training the subject demonstrated transitive inference, consistently selecting D rather than B on the critical B<D test pair. Having demonstrated transitive inference responding on B<D, Gillan extended the series by training a further pair E-F+, and again tested for transitive inference formation using two new non-adjacent test pairs BE and CE (Experiment 1B). Sadie responded correctly on these new test items choosing E over B and E over C.

In both these transitive inference tests the training stimuli formed a linear series. Gillan's next manipulation was to introduce a training pair which disrupted the order of this series by effectively removing the end points. This time Sadie received training with the adjacent pairs, A-B+, B-C+, C-D+, D-E+, E-F+ as had been trained in the previous manipulations, and also with a new pair F-A+ which removed the high and low end points of the series, effectively making it circular. Sadie then received testing on the non adjacent pairs BD, BE and CE which had been tested in the previous two manipulations (Experiment 2A). When the F-A+ pair was introduced, Sadie was much slower to reach criterion than on any of the previous pairs. The introduction of this pair also affected performance on the A-B+ and E-F+ pairs. However Sadie did eventually reach criterion on all six adjacent pairs. When tested, Sadie maintained accurate performance on a high percentage of the directly trained adjacent pairs and this performance was comparable to the

performance shown in the two previous manipulations. However, performance on the non-adjacent test pairs was very different to that shown previously and was close to chance level.

To ensure that disruption of performance was not just due to the addition of a sixth training pair, but was in fact due to that training pair violating the established order of the series, Gillan carried out a further manipulation. Gillan now trained the pairs A-B+, B-C+, C-D+, D-E+, E-F+ and A-F+. This meant that once again six adjacent pairs were trained, but where previously F-A+ had violated the order of the series, this time A-F+ preserved the series order (Experiment 2B). Sadie reached criterion on this task which restored the series order more quickly than she had on the previous task which disrupted the order. When tested Sadie again maintained accurate performance on the adjacent pairs. When tested on the non-adjacent pairs BD, BE, CE, she consistently selected D and E over B and C. This non-adjacent test pair performance was significantly different from performance on the previous manipulation which removed the end points of the series but was not significantly different to performance on the two previous manipulations which maintained the series order. This suggests that Sadie's chance level performance on the non-adjacent test pairs in experiment 2A was not caused simply by the addition of a sixth training pair but was specifically related to the way that pair violated the established linear order of the sequence. This seems particularly likely as, when the training pair which violated the linear order was replaced by one which restored that order, Sadie's performance improved to the same standard as she had shown on the four and five item series.

Bryant and Trabasso (1971) had suggested that accurate performance on the adjacent training pairs was necessary in order for the subject to demonstrate transitive inference. Gillan (1981) now suggested that there were in fact two necessary conditions if the subject was going to demonstrate accurate transitive inference performance; firstly, accurate performance on the adjacent training pairs, and secondly, the stimuli from the series must be ordered on some unidimensional scale. As a result of these experiments, Gillan suggested that transitive inference is a fundamental reasoning process in primates, and this was supported by McGonigle and Chalmer's (1977) findings with squirrel monkeys.

One important aspect of Gillan's (1981) experiment is that the pair comparisons were based on relative amounts of food. Bryant and Trabasso (1971) presented their comparisons in terms of relative lengths of rods, and McGonigle and

Chalmers (1977) used relative weights. Both these presentations used some concrete dimension which may have made it easier to order the stimuli. It was possible that Gillan's presentation using relative amounts of food would be more abstract and would be treated as a series of overlapping discrimination problems rather than forming a transitive series. However, the consistent choice of D and E on the non-adjacent pair tests and the fact that Sadie's performance was susceptible to manipulation of the series order, suggests that the adjacent pairs were in fact ordered into a transitive series. "The results of the present experiments suggest that she had acquired a mental representation of the order of the stimuli in the series" (Gillan, 1981, p.161).

One interesting feature that Gillan (1981) pointed out was that at the time of his and McGonigle and Chalmers' (1977) transitive inference experiments with non humans, two prominent theories of transitive inference performance with humans proposed that language was central to this type of reasoning (Clark 1969; Sternberg 1980). Neither of these language orientated theories explained the data obtained with non-human subjects, and the data from these subjects suggested that language is not a necessary condition for the derivation of transitive inferences. Part of the problem of investigating transitive inference with non-humans had been finding an appropriate way to present the premises. This was made possible by adapting Bryant and Trabasso's (1971) procedure. It has also been difficult to demonstrate stimulus equivalence with non-humans. Possibly this is not because the ability to derive stimulus equivalence is linguistically based but because it proved more difficult than expected to develop ways of presenting the stimulus equivalence task non-linguistically.

Two further studies have repeated Gillan's (1981) procedure of disrupting the established transitive inference sequence. These studies have extended the analysis of transitive inference to non-primate subjects; Fersen, Wynne, Delius and Staddon (1991) - pigeons, Davis (1992) - rats.

Fersen, Wynne, Delius and Staddon (1991)

Fersen, Wynne, Delius and Staddon (1991) suggested that transitive inference testing might be assessing an ability with practical relevance; "the ability to rank objects on a hedonic scale and make judgements about the desirability of items that have never been encountered together must often have adaptive value" (p.334). Fersen et al aimed to show that the ability to form transitive inferences was not limited to primates by demonstrating transitive inferences with pigeon subjects. In

their first experiment they established the series $A > B > C > D > E$ and tested for the emergence of $B > D$. Four out of six subjects completed training and all four subjects showed correct selection of B over D at significantly above chance level. In experiment 2, the five term series was extended to a seven term series by the addition of two new pairs of stimuli, one at either end of the series, $X + A -$ and $E + F -$. This gave six possible non-adjacent test pairs that did not include end items, AC, AD, AE, BD, BE and CE, all of which were tested. The same four subjects that completed experiment 1 were used in experiment 2. Performance for each subject on each test pair was consistent with the derivation of transitive inferences.

In experiment 3 this seven term linear series was closed into a loop by training a new pair $F + X -$. With Gillan's (1981) chimpanzee subject this manipulation had disrupted the transitive inference series and produced random responding on the non-adjacent test pairs. A very similar performance was shown with the pigeon subjects. Following the introduction of the inconsistent item, one subject ceased to respond. Of the three remaining subjects, only two discriminated above chance level on the training pairs and were given the test items. Tests were given on seven non-adjacent pairs, XB, XC, XD, XE, BD, BE and BF. Of the two subjects tested, neither showed stimulus preferences above chance level on any pair. These experiments showed that pigeons are capable of behaving according to transitive inference rules on five and seven term linear series but not on a circular series, performances very similar to that shown by a chimpanzee subject by Gillan (1981).

Davis (1992)

Davis (1992) also investigated transitive inference performances by non-primates in a similar manner, using rats as subjects. Davis established a series $A < B < C < D < E$ and tested for preference for D over B on the non-adjacent BD test pair. Three out of four subjects reached criterion on the training pairs and were tested on BD. Each of these subjects showed highly significant transitive inference performance. A second experiment, training the reversed series $A > B > C > D > E$ and testing for preference for B over D established that this transitive inference performance was not an artifact of how recently the correct premise was reinforced during training.

A third experiment then extended the original series by training $E < F$, which was consistent with the relations already trained, and then made the series circular by training $F < A$, which was inconsistent with the series established. This procedure replicated that carried out by Gillan (1981). Davis (1992) stated that while it was clear that disrupting the internal transitive relations should affect the way in which

inferences were made on the test pair BD, it was not clear how this disruption would manifest itself. Should it produce random performance on BD, or would it cause the series to reverse, giving a preference for B over D? Also, would the introduction of this inconsistent relation cause immediate disruption of the internal structure or would the disruption develop gradually? Davis counter-balanced for the order in which the disruptive training occurred. Three rats who had already completed experiment 1 and had received training on the series $A < B < C < D < E$ now received training on the relations $E < F$ and $F < A$. Three experimentally naive rats were first taught $E < F$ and $F < A$ before receiving training on the series $A < B < C < D < E$. After training, all subjects were tested on two non-adjacent pairs, BD and BE. All subjects were affected by exposure to information that altered the logical structure of the transitive series. Although there were no significant differences between the two training conditions on the non-adjacent pairs, Davis believed there were some subtle effects according to the order in which premise inconsistencies were introduced. Subjects exposed to the inconsistencies before training seemed to have greater difficulty reaching criterion on the training trials. When tested, naive subjects failed to show a preference for stimulus D and in the subjects who had previously demonstrated transitive inference this preference for D was eliminated or reversed. On the test pair BE, E had formerly been the endpoint of the series and so it was expected that on the consistent series there would be a very strong preference for this stimulus. On the inconsistent series this preference for E was eliminated in 2 out of 3 subjects who had previously demonstrated transitive inference. The naive subjects showed a slight (non-significant) preference for E on the inconsistent series but Davis believed that this could be accounted for by reinforcement of E in the final stages of training with these subjects.

These results with rats as subjects were very similar to the results obtained with other non-human subjects. Both human and non-human subjects seem to be capable of making transitive inferences, as long as the stimuli form an ordered linear sequence. When the series' linear order is disrupted, transitive inference responding is also disrupted. However, Davis (1992) points out that while there may be evidence for a continuum of mental abilities, it may not be wise to assume that different species approach transitive inference problems in the same manner. As there are considerable neurological differences between humans, chimpanzees, monkeys, rats and pigeons it would hardly be surprising to find considerable cognitive differences as well.

"If different species are truly approaching logical transitivity in the same way , then, to put the case crudely, either rats must be stretching their ability to the breaking point or humans must be barely using theirs" (Davis, 1992, p.348)

So while non-humans may be able to make transitive inferences it may not be wise to assume that they do this in the same way or with the same ease as human subjects. It is also doubtful that they can solve problems of the same complexity as those which human subjects can deal with.

Careful experimental investigation has demonstrated that non-human subjects are capable of solving quite complex cognitive tasks such as transitive inference and probably, stimulus equivalence. However, their behaviour on these tasks may not be the same as that of human subjects, so while Thorndike's view that data from non-human subjects is relevant to tracing "the origin of human faculty", it is important to be cautious when comparing human and non-human performance on similar tasks. A detailed analysis of both human and non-human responding is necessary in order to ascertain if the processes operating are similar.

CHAPTER 4

Examples of Human Cognition

Chapter 3 considered the demonstration of cognitive abilities in non-human subjects. This chapter looks at the demonstration of the same abilities in human subjects.

Transitive Inference

To make a transitive inference requires the co-ordination of two pieces of information to reach a correct inferential conclusion. For instance, if stimuli A and B are related to each other by relation r (ArB), and stimuli B and C are related to each other by the same relation r (BrC), then it can be inferred that A and C must also bear relation r to each other (ArC).

Piaget (1928) frequently used transitive inference tasks as a measure of cognitive development in children. Piaget, with others, proposed that children were unable to make these inferences until they reached the stage of concrete operations at around age seven. Piaget felt that pre-operational children were unable to make transitive inferences because they were dominated by immediate perceptual input and thus unable to reorganise this input to combine the information from two relations. Transitive inferences became possible when the child progressed to a "*relativistic*" conception of relations (Breslow 1981) and understood that a term can have two relations simultaneously, e.g. B can be smaller than A ($B < A$) and larger than C ($B > C$). This allows the child to co-ordinate the two independent relations $B < A$, $B > C$ to infer the relation $A > C$.

Bryant and Trabasso (1971)

Some researchers, however, have challenged Piaget's conclusions that children can not make inferences before the age of seven. Bryant and Trabasso (1971) suggested that much younger children were able to make transitive inferences provided they could remember the comparisons they had to combine. They felt that previous failures may have been failures of memory rather than failures of inferential ability. Bryant and Trabasso controlled for this possibility by training the initial comparisons very thoroughly, and also checking retention of these comparisons during testing. They tested for transitive inference by teaching four overlapping pair comparisons $A > B$, $B > C$, $C > D$, $D > E$ and testing performance on the non-adjacent pair $B > D$, using a series of wooden rods of differing lengths and colours.

This sequence of five stimuli, rather than the more usual three stimuli, was designed as a control against "absolute" responding. Bryant and Trabasso pointed out that when three terms are used the correct response to the extreme properties was the same for both the initial comparisons and the test for inference. For example, given the relations $A > B$, $B > C$, A is always "larger" ($A > B$ -initial comparison, $A > C$ -inferential comparison) and C is always "smaller" ($B > C$ -initial comparison, $A > C$ -inferential comparison). Therefore, a subject could respond appropriately on the $A > C$ test but without using an inferential strategy. This could be done by parroting the verbal labels "larger" and "smaller" from the initial training, or simply always selecting A and never selecting C. The five stimulus sequence used by Bryant and Trabasso controls for this possibility. The training given establishes the series $A > B > C > D > E$ so that the correct response on the transitive inference test is $B > D$. During training the stimuli on the transitive inference test have both sometimes been "larger" ($B > C$, $D > E$), and sometimes "smaller" ($A > B$, $C > D$). In this case correct responding requires co-ordination of the training relations via the middle term C.

Bryant and Trabasso (1971) trained this sequence using five coloured rods of different lengths and colours. These were presented to the subject embedded in a container so that each rod protruded from the top of the container by one inch. This was designed to force the subjects to use the different colours when making a choice between different lengths. The training and test procedures were carried out with three groups of children, four year olds, five year olds and six year olds. All three age groups showed good evidence of deriving transitive inferences. This occurred whether the subjects received visual feedback during training and were shown the lengths of the rods after each trial, or whether they received verbal feedback and were just told which rod was longest and which shortest when they had made their selection. However, Bryant and Trabasso noted that while performance on the transitive inference task was significantly above chance level for all age groups, in no case was it perfect. Analysis of the data showed that the probability of making a correct inference on the BD task was directly related to the probability of jointly recalling the information from the necessary training pairs BC and CD. This seemed to support Bryant and Trabasso's hypothesis that children younger than seven could make transitive inferences provided they could remember the necessary comparisons.

Lutkus and Trabasso (1974)

Lutkus and Trabasso (1974) replicated the procedures used by Bryant and Trabasso (1971) with pre-operational adolescents with learning disabilities. Again, a major concern was the age at which different subjects can make transitive inferences. Piaget (1970), Smedslund (1963, 1965) and Youniss and Murray (1970) had all suggested that normally able children could not reliably make transitive inferences until the age of about 7 to 9 years, while McManis (1969, 1970) had suggested that learning disabled children were even further delayed, until a mental age of 10 to 11 years. Braine (1959, 1964) and Bryant and Trabasso (1971) had found evidence of transitive inferences with normally able children as young as 4.

Lutkus and Trabasso (1974) considered what were the necessary criteria for deciding whether or not a person was capable of making transitive inferences. Smedslund (1969) suggested that there were two kinds of errors that might be made when deciding if a person was capable of making inferences. The first is a "false-negative" error, where a subject might be classed as non-transitive when in fact that subject is capable of making transitive inferences. This sort of error might occur if the subject does not understand the information given, or the instructions, or may forget either the information or the instructions. The second kind of error is a "false-positive" when the subject is classed as able to make a transitive inference when the subject does not in fact have this ability. This might occur when a non-inferential solution is available for an inference problem, or even if the subject just guesses correctly.

Piaget and others had tried to guard against false-positives by requiring a verbal explanation of the subjects solution to the problem. It was unlikely that a subject would use a non-transitive solution to solve the problem while also explaining the correct transitive solution. However as Thayer and Collyer (1978) pointed out "A child may be able to make the correct judgement on a transitive inference task but may be unable to verbalise an adequate explanation for the choice" (p.1334). Therefore, although the child may have used an inferential solution to the problem, a failure to verbalise this solution might result in a false negative assessment of that subjects inferential ability. In contrast Bryant and Trabasso's (1971) experiment used a "symptom-response" to define the presence or absence of transitive inference. Bryant and Trabasso had trained four overlapping pair discriminations to establish the series $A > B > C > D > E$. The subjects were then tested for performance on the BD pair of $B > D$. Success on the BD pair was the "symptom-response" and the experiment was designed so that success on this pair could only be explained

by transitive inference. Using this procedure Bryant and Trabasso had suggested that children as young as four could make transitive inferences. Lutkus and Trabasso (1974) aimed to replicate this procedure with adolescents with learning disabilities to see if these subjects would show transitive inferences at mental ages comparable to normally able subjects.

Lutkus and Trabasso (1974) used Bryant and Trabasso's (1971) procedure to test for transitive inference with 40 adolescents with learning disabilities. These subjects were divided into two groups of 20 subjects, one group with a mental age (MA) of 5.6 years, and one group with a mental age of 6.4 years. These subjects received training on the four comparisons $A > B$, $B > C$, $C > D$, $D > E$. In the test phase the subjects were tested on the ten possible stimulus pairs - the four training pairs AB, BC, CD, DE and six new non-adjacent transitive pairs, AC, AD, AE, BD, BE, and CE. This included the test pair BD, the only test item that did not include one of the stimuli from the extreme ends of the series, and thus the most stringent test for transitive inference. These subjects learned the four initial direct comparisons very quickly and performance on this task was not significantly different to performance by normally able children of comparable mental ages. On the second training phase there were no significant differences between the two adolescent MA groups, but both groups took over five times longer to reach criterion than normally able children of the same MA. When given the test items, the learning disabled subjects scored significantly above chance level on all the transitivity test pairs. This is particularly important on the "double reversible" BD pair as high proportions of correct choices on this pair reflects inferential behaviour by the subject.

Both learning disabled MA groups performed like the normally able children on the test pairs involving the end items A or E. However, performance on the comparisons BC and CD and on the transitive inference pair BD was lower for the learning disabled subjects. This performance on BD was shown to be related to the retention of the initial BC and CD comparisons, thus supporting Bryant and Trabasso's (1971) hypothesis that transitive inference test performance was directly related to retention of the initial premises. The distribution of performance on the BD test pair was compared for the learning disabled MA groups and the normally able comparable MA groups. There was no significant difference in distribution for either MA group or its comparable normally able group.

This experiment showed that learning disabled subjects with mental ages that would be regarded as "pre-operational" were able to make transitive inferences, and their success rates on this task were similar to normally able children of comparable mental ages who had received the same training and testing. Lutkus and Trabasso also concluded that verbal explanations of the subjects behaviour were not necessary for concluding that there had been transitive reasoning on this task. Rather, they suggest that verbal explanations may produce MA differences in findings as linguistic ability is age dependent. They suggest that explicit transitivity can be verbalised by normally able children at around eight years old, but that these data showed that children or learning disabled adolescents with mental ages younger than eight can make use of relational information in a transitive way - even if they can't always verbalise the steps they go through.

Kuczaj and Donaldson (1982)

Kuczaj and Donaldson (1982) suggested that there may be a distinctive developmental pattern in the way children learn to make transitive inferences. They felt that young children's ability to employ language and to create concepts suggests that children as young as four do indeed possess the cognitive sophistication necessary for inferential activity. However, they also suggested that while young children may be able to make transitive inferences they may not always make them appropriately and this ability must develop over time. Kuczaj and Donaldson suggested a pattern of four stages before children were able to make appropriate, mature transitive inferences. Firstly, an over specific phase when the child understands only a few terms that may form transitive relations e.g. bigger/smaller. This is followed by a gradual increase of the relational bases that may enter into transitive relations. This then results in overgeneralisation of transitive inference to include inappropriate terms e.g. love/kick that do not necessitate transitive relations. Finally these overgeneralisation errors are gradually eliminated to result in appropriate application of transitive inference skills.

Kuczaj and Donaldson ran two experiments to test this. They found that four and five year olds tended to give overgeneralised transitive responses and correct transitive responses. Six and seven year olds tended to give consistently only correct transitive responses. However, eight year olds showed the same pattern as four and five year olds, tending to give both correct and overgeneralised transitive responses. While this developmental pattern emerged, it failed to reach statistical significance. Kuczaj and Donaldson suggested that the responses of the four and five year olds may have been based on a response strategy as correct transitive

responses were rarely seen without overgeneralised transitive responses; this was tested in the second experiment. Experiment 2 replicated the procedure in experiment 1 but also tested if the subjects were likely to use a response strategy when responding. The data from this experiment suggested that the four year old children's overgeneralisations were indeed due to a response strategy as the children who gave correct transitive responses and overgeneralised transitive responses also gave transitive responses in the strategy task. However, this response strategy was not shown by the older children who gave correct transitive responses and overgeneralised transitive responses. This suggested that the overgeneralisation errors of the eight year olds may have been true overgeneralisation errors.

Kuczaj and Donaldson believed that these results demonstrated that transitive inference skills followed a specific to general developmental pattern, with children initially applying their transitive inference skills to a small number of appropriate terms followed by a gradual broadening of the application of these skills. This broadening could result in overgeneralisation on the transitive skills to include inappropriate relational terms. The child must then learn to exclude these inappropriate terms from their repertoires while still continuing to add new, appropriate terms.

Artman and Cahan (1993)

A study by Artman and Cahan (1993) suggested that one of the most important influences on the development of transitive inference skills was a child's experience of schooling. Artman and Cahan felt that the transitive inference literature had tended to focus on two issues: the age at which children are first able to make transitive inferences, and how children of different ages actually make inferences. "However, little attention has been devoted to the causal model underlying this developmental process" (p.753). The authors suggested four reasons why schooling may be relevant to the development of transitive inference. First, schooling was explicitly aimed at developing intellectual abilities. Secondly, Artman and Cahan suggested there are structural similarities between cognitive tasks and common school assignments:

"(a) Both are self-contained: (b) there is typically one "correct" answer: (c) the content of the problems are often of limited, academic interest: and (d) they are solved for their own sake" (p.753)

Thirdly schooling may indirectly affect cognitive development as the higher the grade level, the greater the cognitive demands/expectations and the richer the test

taking experience. Fourthly, they felt that the empirical evidence pointed to an effect of schooling on the development of cognitive skills (Ceci 1991).

Artman and Cahan also believed that schooling could have a positive effect superficially on transitive inference skills. "The school experience offers many opportunities to compare objects on various unidimensional scales, such as physical or psychological traits, all of which are characterised by the transitivity of the order relation" (p.753). Bryant and Trabasso (1971) had already noted that if a child was unable to co-ordinate two items of relational information in order to make a transitive inference then that child would have difficulty understanding the most elementary principles of measurement. It would seem reasonable then that increased experience with measurements and relative values would also help to develop transitive inference skills.

However, some theoretical considerations suggest that the effect of schooling on transitive inference ability would be lower than the effect on other types of cognitive performance. An intelligence development approach predicts that schooling will only affect the development of "crystallised" abilities and "fluid" abilities such as transitive inference are influenced by "incidental" childhood learning. Also, neo-Piagetian approaches which regard working memory functions as essential to successful performance on transitive inference tasks would attribute development to biological maturation rather than external interventions such as schooling.

All the same, Artman and Cahan believed it could still safely be assumed that schooling affected cognitive development and that the issue was the size of this effect; both the absolute size of the schooling effect and its size relative to the effect of chronological age. They investigated this by giving a transitive syllogism test of 22 three-term series problems to pupils in the fourth, fifth and sixth grades at school. They then analysed the data from this using a between-grades regression discontinuity paradigm to carry out post-hoc estimations of schooling versus age effects on success in the test. In this design the estimated effect of a one year difference in chronological age in a given grade equals the difference between the oldest and youngest students in that grade in mean predicted scores, and the estimated effect of one year schooling equals the difference in mean predicted scores between the youngest children in one grade and the oldest children in the lower adjacent grade (Artman and Cahan 1993, Cahan and Cohen 1989, Cahan and Davis 1987).

Analysis of the data in this way indicated an effect of schooling on test scores. Furthermore, the estimated effect of one year of schooling and one year of age on the total score showed that schooling is a major influence on the increase of transitive inference test scores as a function of age. Artman and Cahan concluded that "schooling affects the development of transitive inferences independently of chronological age and associated factors" (p.757). Two reasons they suggested might underlie this effect were, first that there might be a direct effect of schooling by providing experience with order relations and opportunities to practice relational reasoning. Secondly, schooling may develop other relevant skills such as memory, seriation and test experiences.

Siemann and Delius (1993)

Siemann and Delius (1993) examined transitive inference in the context of implicit/explicit cognition. There was some evidence that "humans are . . . capable of complex perceptual and mnemonic feats without necessarily being aware of it" (p.364). Siemann and Delius felt that transitive inference tasks provided a simple way to examine implicit/explicit deductive cognition. In doing this they used the training paradigm devised by Bryant and Trabasso (1971) and extensively used with non-human subjects. Twenty-four normally able adult subjects (mean age 23) were taught the overlapping pairs A+B-, B+C-, C+D-, D+E-, and E+F- (in each pair, + denotes the reinforced stimulus, and - the unreinforced stimulus) to establish the series A>B>C>D>E>F. This was done in the context of a computer game where the symbols comprising each stimulus were presented on doors. Selecting the correct symbol on the door produced a reward of one coin for the cartoon character and selecting the wrong door produced a cost of one coin for the character. Three transitive inference test pairs were presented, BD, CE and BE, and these were also tested in the context of the computer game (but without reinforcement). After the computer game section of the experiment was completed, the subjects were asked to complete a questionnaire asking what they thought the purpose of the experiment was and what, if any, strategies they had used. The subjects were also given cards of each of the six stimuli and asked to arrange them into any orderly arrangement that their experiences suggested to them. Of the 24 subjects, 15 "solved" the problem and responded at significantly above chance level on the test pairs. Of these 15 subjects, the questionnaires showed that only 8 subjects recognised the stimulus hierarchies. These "aware" subjects had verbalised the stimulus relation and were able to use this information to respond on the test pair. The other seven subjects who solved the problem did not know what the

computer game was about. When questioned they said they had either guessed, or could not explain their responding, or used strategies that were not in fact correct. However this group's performance was no different in speed or accuracy to the eight "aware" subjects.

The seven "unaware" subjects were initially puzzled by the card organising task, with several insisting that they could not do it. However, five of these subjects did eventually produce the correct sequences when pressed to complete the task. Of the nine subjects who did not solve the transitive inference task, only one was aware of the nature of the task, and none succeeded on the card-organising task.

These results showed that humans were able to respond according to deductive rules without being conscious of those rules, and importantly, the performance of these subjects who were responding "implicitly" was not significantly different to the performance of subjects who responded "explicitly". Siemann and Delius suggested that this might be evidence that even subjects who respond explicitly "may be behaving transitively without recourse to the propositional rules of classical logic" and that as such "transitive inference competence may be an evolutionary ancient trait" (p.366) , and therefore can be demonstrated in a number of non-human subjects.

Siemann and Delius (1994)

A further study by Siemann and Delius (1994) pointed out that one of the values of the procedure of training adjacent pairs $A+B-$, $B+C-$ etc. was that it demonstrated the economy of a single hierarchical ordering. This economy of effort might have adaptive value for many organisms; "As information processing is associated with definite costs, any reduction of redundancy within this incoming stream generally signifies an advantage in fitness" (p.531). As previous research had shown, training four overlapping pair relations $A>B$, $B>C$, $C>D$, $D>E$ was sufficient to specify the series $A>B>C>D>E$, which in turn allowed inference of six further nonadjacent pair relations. If the number of original pair comparisons increased, the importance of having an economical way of organising these relations would also increase.

This study extended previous work with non-human subjects (Davis 1992, Fersen et al 1991, Gillan 1981) which established a linear transitive sequence and then made that series circular. In this experiment Siemann and Delius used overlapping pair comparisons to train stimulus sequences with three groups of subjects. All subjects were taught 10 pair comparisons. For group G0 all 10 comparisons

complied with a linear structure $A > B > C > D > E$. For group G1, one of these pairs, $E + A -$, was inconsistent with this linear structure, while for group G3, three of these pairs were inconsistent with this linear order, $D + A -$, $E + A -$ and $E + B -$ (see Table 4:1). Thus Siemann and Delius intended that this training would establish a linear series with group G0, a circular series with group G3, and what they referred to as an "odd" series with group G1 (see Fig. 4:1). Siemann and Delius carried out this training procedure with both human and pigeon subjects, and examined the differences in performance between these subject groups and between the G0, G1 and G3 groups on the different tasks.

Thirty normally-able human subjects (mean age 22.8 years) took part in this experiment, 10 in each group. The task was presented on a computer screen, with each pair of stimuli (simple geometric figures) appearing side by side. After training and testing by computer, the subjects were given a questionnaire and asked what they thought the experiment was about. Finally they were given five cards, one of each of the stimuli, and asked to arrange the cards into any sequence their recent experience suggested to them. Additional stimulus cards were available if the subject needed them.

There were significant differences in the number of learning blocks required by each group to reach criterion. Group G0 subjects required a mean of 3 blocks, G1 subjects a mean of 5.1 blocks and G3 subjects a mean of 6.3 blocks. Two subjects in group G3 failed to achieve the accuracy criterion for testing. The data were analysed according to the symbolic distance between the stimuli. For instance, the adjacent pairs AB, BC, CD and DE had a symbolic distance of 1 (d_1). Non-adjacent pairs AC, BD, and CE had symbolic distance d_2 , AD and BE had symbolic distance d_3 and AE symbolic distance d_4 . This way of describing the distance between the stimuli in each pair was based on the linear ordering of the series and thus not entirely appropriate for the sequences established in G1 and G3, but it did give some indication of relative performance on these tasks compared to a standard transitive inference task. The overall choice accuracies of the groups did not show any significant differences, but as all three groups had been trained to the same accuracy criterion this is not surprising. However the overall choice latencies were highly significant, increasing from G0 to G3. There were also significantly different choice latencies for the different symbolic distances. Group G0 showed a symbolic distance effect, with choice latencies decreasing as the distance separating the stimuli increased. Group G3 also showed this symbolic

distance effect, an unexpected result. The G1 data showed no significant latency differences.

When given the questionnaire and tested on the card sorting task, 9/10 G0 subjects reported memorising a serial order and referring back to that when tested. These subjects also laid out the cards in the correct order. In group G1, six subjects reported using a serial order strategy. Of these six subjects, four were able to lay the cards out in the correct sequence and three of those required an extra A stimulus card to complete the inconsistent E>A pair. However, of the G1 subjects who did not report using any strategy and who did not sort the cards correctly, two subjects still responded at above chance level in both speed and accuracy. This replicated Siemann and Delius' (1993) finding of a dissociation between knowledge and performance. The G3 subjects tended to report memorising the stimuli in pairs using verbal markers. None of the G3 subjects sorted the cards into the correct overall structure.

Nine pigeon subjects also received training on this experiment, three in each group. The stimuli were presented on two horizontal side by side pecking keys. All three pigeons in groups G0 and G1 and one pigeon in group G3 completed training. The remaining two pigeons in group G3 ceased to respond after 3 and 4 sessions respectively. Subjects in group G1 and G3 showed considerably slower learning than subjects in G0. The mean error rate of group G0 was significantly below chance level. When performance (% errors) was assessed across the symbolic distances, group G0 showed a clear symbolic distance effect. Group G1 showed considerably increased error rates for D4 compared to the other distances. This result is not surprising as only the D4 pair AE was inconsistent with a linear series. The remaining subject in group G3 showed nearly random choices over all distances.

In summary then, both human and pigeon subjects learned and retrieved the pair comparisons in this task most efficiently when the stimuli were organised in a linear hierarchical order. As the number of exceptions to this order increased, performance by both groups of subjects declined. However, the human subjects were all able to cope with even the most difficult task, G3, albeit at the expense of long reaction times. However, the pigeon subjects in the G3 task either ceased to respond or responded at around chance level. The questionnaire and card sorting task showed that for the human subjects, perception of the structure of the task was affected by increasing departures from a linear structure. Siemann and Delius

suggested that these results showed that a hierarchical linear network had advantages for the processing of information as it allowed for economy of effort in coding and retrieving the information.

Stimulus Equivalence

Many of the early experiments developing the stimulus equivalence paradigm (e.g. Sidman 1971, Sidman and Cresson 1973, Spradlin, Cotter and Baxley 1973, Dixon and Spradlin 1976) used people with learning disabilities as subjects. Data from these experiments showed that even subjects with quite severe learning disabilities could produce the emergent relations which define equivalence relations (Sidman and Cresson 1973). The paper by Sidman and Tailby (1982) defining the properties of equivalence relations, reflexivity, symmetry and transitivity, clearly demonstrated the emergence of stimulus relations with normally able children. Other papers examined the possibility of demonstrating equivalence relations with non-human subjects (e.g. Sidman et al 1982, Dugdale and Lowe 1990, D'Amato et al 1985). These studies tended to conclude that it was difficult, if at all possible to unequivocally demonstrate stimulus equivalence with non-humans. So given that it was possible to demonstrate equivalence with human subjects with quite limited abilities, but not to demonstrate equivalence with even the higher primates, it seemed reasonable to explore the limits to which human subjects demonstrated equivalence relations.

Devany, Hayes and Nelson (1986)

One study which tried to determine which groups of human subjects would or would not demonstrate the emergence of equivalence relations was carried out by Devany, Hayes and Nelson (1986). Devany et al suggested that stimulus equivalence appeared to relate well to the issue of symbolic activity in that a "symbol" and its "referents" form a class of mutually substitutable elements.

"The relations among the members of an equivalence class appear to approximate what psycholinguists and others mean when they say that a word represents or "stands for" its referent in a way that a conditionally related response does not". (p.244)

Just as the relations in an equivalence class are not uni-directional (like a conditional relation) but rather functionally reversible, so the relation between a symbol and its referent must be bi-directional. "A word "stands for" another event only if the event "is called" the word" (p.244). Devany et al suggested that if equivalence classes are in fact connected to symbolic/verbal activity then it would

be expected that equivalence relations would be easy to demonstrate with human subjects but difficult to demonstrate with non-humans, as appeared to be the case; certainly in equivalence research at that time. Therefore Devany et al speculated that the ability to form equivalence classes is related to language acquisition or use, and to test this compared the performance of normally developing children on an equivalence test with the performance of language impaired children on the same test.

Three groups of children took part in this experiment (four children in each group): normally able children, learning disabled children who demonstrated expressive speech or sign skills, and learning disabled children who lacked functional speech or signing. Each child was matched with a child in each of the other two groups on a measure of mental age. There were four levels of mental age for each group, ranging from 14 to 36 months. The children were taught four conditional discriminations, A-B, A-C, D-E, D-F, to try and establish two equivalence classes; ABC and DEF. The stimuli were made up of animal like figures.

The learning disabled subjects with no expressive language required significantly more trials to complete training than either of the other two groups. There were no significant differences between the normally able subjects and the learning disabled subjects with expressive language in the number of trials needed to complete training. The learning disabled subjects with no expressive language also required more prompts (visual and manual) in training

When given the equivalence test, all the subjects in both the normally able group and learning disabled with expressive speech group showed high percentages of correct responding, indicating that they had derived the predicted equivalence classes. Responding by the learning disabled subjects with no expressive language remained near chance level throughout, indicating that they had not derived the equivalence classes. These results were consistent across children with varying mental ages in each group. Statistical analysis showed that on the equivalence test, there was no significant difference between the performances of the normally able subjects and the learning disabled subjects with expressive language. However there was a significant difference between these two groups performances and performances by the learning disabled subjects without expressive language., with the two "language-able" groups performing significantly better on the equivalence test.

These results identified a particular group of subjects (learning disabled with no expressive language) who failed to show equivalence class formation, even though these subjects had each been matched on a measure of mental age with subjects in other groups who did show equivalence class formation. This failure could not be explained by a failure to learn conditional discriminations as all the subjects reached criterion on the training tasks. Although the learning disabled subjects without expressive language required substantially more training to reach this criterion. Devany et al felt that the data provided support for the view that symbol use/language and the ability to derive equivalence classes are closely related. The results also included data from the youngest subject reported, to that point in time, to show equivalence class formation; 25 months (normally able group).

Devany et al did note however that caution was needed when interpreting these data. For instance, the learning disabled children differed from the other children in ways other than in a failure to speak or sign. They required substantially more conditional discrimination training and more prompting within that training. Also, using mental age scores provided only a very crude way of matching the children in the different groups. For a child without language skills to be matched by mental age to a child with language skills, the first child would have to have better developed non-verbal skills. So the matching of subjects and assignment to subject groups may not have been perfect. Devany et al also noted that while the data seemed to support the view that symbol use and language was related to equivalence class formation, these results were correlational and "it is not possible to say . . . whether the ability to form equivalence classes is a precursor of symbol use, a product of it, or if the two are both a reflection of the same process" (p.254). However the data did seem to provide some support for Devany et al's hypothesis that language-ability and equivalence class formation were related. The experiment also seemed to identify a group of human subjects who did not readily demonstrate equivalence relations, despite having learned the requisite conditional discriminations.

Barnes, McCullagh and Keenan (1990)

One of the problems with interpreting the data from Devany, Hayes and Nelson (1986) was determining to what extent the subjects' expressive language abilities may have affected their performance on the equivalence task, and to what extent these results were confounded by the subjects' degree of learning disability. A study by Barnes, McCullagh and Keenan (1990) replicated the procedures used by Devany et al with three groups of subjects who differed in their degree of

expressive language ability but none of whom had learning disabilities. In this experiment the three groups of subjects were: normally developing pre-schoolers, normally developing, partially hearing children who demonstrated some expressive language, and normally developing, partially hearing children with little or no expressive language. There were two subjects in each group. As Devany et al had reported that the youngest human subject they had found to demonstrate stimulus equivalence was aged 2 years and 1 month, the partially hearing children were classified into Set 2 (expressive language age of 2 years and above) and Set 3 (expressive language age of 2 years and below). Set 1 were the non-hearing impaired children. The expressive language ages of the subjects had been established using the Reynal Developmental Language Scales.

The subjects were taught four conditional discriminations, A1-B1, A2-B2, A1-C1 and A2-C2. They were then tested for the emergence of the B1-C1 and B2-C2 equivalence relations. When tested, both the children in Set 1 (normally hearing) and both the children in Set 2 (partially hearing, expressive language age above 2 years) clearly demonstrated the emergence of equivalence classes (95% - 97.5% correct). Of the two children in Set 3 (partially hearing, expressive language age of below 2 years), one subject (Clare) clearly demonstrated equivalence relations (97.5% correct). The other subject (Claudia) initially scored 0% correct when tested. Following this unusual result, this subject received training with a different set of stimuli and was re-tested. On the second test the subject scored 50% correct, exactly chance level, demonstrating that equivalence classes had not been formed.

These results replicated the findings of Devany et al (1986) in that all the subjects in Sets 1 and 2 (who had an expressive language age of above 2 years) demonstrated the formation of equivalence classes. Of the subjects in Set 3 (expressive language age of below 2 years) one subject did demonstrate equivalence relations and the other did not. The subject who did demonstrate equivalence relations was found to have an expressive language age of 1.5 to 2.0 years, which would put her near the language ability of 2 years which Barnes et al felt might be critical for the development of equivalence relations. The other subject, Claudia, had an expressive language age of 1.0 to 1.5 years, the lowest age of all the subjects tested. So although the results did not precisely match Barnes et al's prediction that the subjects in sets 1 and 2 would demonstrate stimulus equivalence while the subjects in set 3 would not, the data still fit quite well with the hypothesis that expressive language skills and the ability to show equivalence relations are in some way related. Equally importantly, in this experiment the

subjects were all normally developing children and thus any differences in performance cannot be attributed to some degree of learning disability.

The subjects in the experiment by Devany et al (1986) who failed to show stimulus equivalence had also required significantly more training trials to reach criterion than subjects who did demonstrate equivalence relations. Devany et al suggested that the discrepancy might be due to some unspecified aspect of the subjects' learning disabilities. This suggestion seems to be supported by Barnes et al's findings, as the subject who failed to demonstrate equivalence, Claudia, actually required fewer training trials to reach criterion than did another subject Aidan (normally hearing) who did demonstrate equivalence (120 vs. 210 trials).

Like Devany et al (1986) the results from the experiment by Barnes et al (1990) seemed to suggest a link between expressive language ability and the ability to derive equivalence relations (to be discussed in greater detail later). However, as the findings from the subjects in Set 3 show, the nature of this relationship was not clear; but both experiments identified a group of human subjects for whom conditional discrimination training did not result in the emergence of equivalence relations, even though the same training produced equivalence relations with other subjects matched on certain measures.

Lipkens, Hayes and Hayes (1993)

A study by Lipkens, Hayes and Hayes (1993) took a different approach to investigating the demonstration of equivalence relations with normally able children. Lipkens et al carried out a longitudinal study of the development of equivalence relations with one child. The child, Charlie, was 16 months old at the beginning of the study and 27 months old at the end. Devany et al and Barnes et al had tried to isolate one aspect of development (expressive language) that they thought might be important in the development of equivalence. In this study, Charlie was tested on a number of measures on several occasions during the study. By this Lipkens et al hoped to establish the stages at which different aspects of derived relations would emerge.

Steven Hayes had proposed a relational frame account of the derived relations found in stimulus equivalence (S.C. Hayes, 1991). While Hayes' relational frame account will be discussed in more detail in a later chapter, Lipkens et al used the terminology of relational frame when discussing their results and these terms should be specified here. The definition of stimulus equivalence by Sidman and

Tailby (1982) requires the demonstration of three properties, reflexivity symmetry and transitivity. Relational frame also has three defining characteristics; mutual entailment, combinatorial entailment and transfer of function. This study by Lipkens et al tested for two of these properties, mutual entailment and combinatorial entailment. In mutual entailment, if A is related to B in a given context as a result of training, then B is related to A in that context by derivation. For instance, if A is better than B, then by derivation B is worse than A. Symmetry would be another example of a relation of mutual entailment. The second characteristic, combinatorial entailment, implies that if A is directly related to B and B is directly related to C, then some relation is entailed between A and C. Transitive and equivalence relations as defined by Sidman and Tailby (1982) would be relations of combinatorial entailment. These relations of entailment as defined by Hayes (1991) encompass a wider range of possible stimulus relations than the relations proposed by Sidman and Tailby. However, the training and testing carried out by Lipkens et al is very similar to that used in standard equivalence experiments. In considering Lipkens et al's procedures in this experiment, tests for mutual entailment were roughly synonymous with tests for symmetry and tests for combinatorial entailment were roughly synonymous with tests for transitivity and equivalence.

Lipkens et al hypothesised that a longitudinal study might be an effective way of examining the processes underlying stimulus equivalence and supporting or refuting some of the theories about the basis of stimulus equivalence.

"If derived relations are a primitive process, no developmental trend need be posited. If they are mediated by language, they should not be evident in very young infants but should emerge as language develops. If they are instances of learned behavior that underlying (sic) language, they should emerge early and should show clear developmental trends" (Lipkens, Hayes and Hayes, 1993, p.204)

Lipkens et al also gave tests for stimulus exclusion. A standard exclusion test might involve presentation of a novel word as sample and a known word and a novel word as comparisons. Research has shown that normal and learning disabled children (e.g. Hutchinson 1986, McIlvane and Stoddard 1981) and animals (Schusterman and Krieger 1984) tend to select novel comparisons given novel samples. Lipkens et al argued that there were two types of exclusion - exclusion with mutual entailment and exclusion without mutual entailment. Lipkens et al suggest that selecting a novel comparison given a novel sample could occur

because of: a mutual relation of difference between the novel sample and the novel comparison, a mutual relation of co-ordination ("sameness") between the known sample and the known comparison, and a derived mutual relation of co-ordination between the novel sample and the novel comparison. In this case, the relation between the novel sample and the novel comparison would be one of exclusion with mutual entailment. However, selecting a novel comparison given a novel sample might be based on direct characteristics of the stimuli such as avoidance of known forms. Lipkens et al term this exclusion without mutual entailment. Hayes and Hayes (1989, 1992) had argued that "behavior based on derived stimulus relations that show properties of mutual and combinatorial entailment is verbal behavior by definition" (Lipkens, Hayes and Hayes, 1993, p.205). Therefore they termed exclusion with mutual entailment "verbal exclusion" and exclusion without mutual entailment "non-verbal exclusion". One of the aims of the study was to see if there was a developmental transition between non-verbal and verbal exclusion.

The subject in this study was Charlie, a normally developing infant. Charlie was aged 16 months 18 days when the study began and 27 months and 11 days when it finished. The experiment used both visual and auditory stimuli. The visual stimuli were large line drawings of familiar objects (e.g. ball, cat) and of novel objects (strange and prehistoric animals). The auditory stimuli were familiar spoken words (e.g. dog, horse) and novel words (e.g. ui, neus).

Experiment 1 tested for relations of mutual entailment. Charlie was taught A-B/picture-name relations (given the picture, what is its name), and then tested for the emergence of B-A/name-picture relations (given its name, where is the picture). The training and test tasks were then reversed with a new set of stimuli so that Charlie was taught B-A (name-picture) relations and tested on A-B (picture-name) relations. Charlie was 16 months and 22 days when picture-name testing began. In picture-name training, Charlie was taught A1-B1, A2-B2 and tested for B1-A1 and B2-A2. Charlie was then taught A3-B3 and A4-B4 and tested on B3-A3, B4-A4. In training trials Charlie did not always produce the stimulus name when asked "What is this?", A1B1 - 50%, A2B2 - 30%, A3B3 - 33%, A4B4 - 83%. However, when tested on the name-picture task, Charlie demonstrated relations of mutual entailment B1A1 - 83%, B2A2 - 100%, B3A3 - 100%, B4A4 - 67%. At the time these tests finished Charlie was 17 months and 7 days old.

Charlie was then taught B-A (name-picture) relations B5-A5 and B6-A6. This set of relations proved much more difficult to teach. Initially training consisted of

presenting the two pictures and asking "Where is OEF/B5 (or UI/B6) ?". During testing one picture was presented and Charlie was asked "What is this?". As this procedure seemed much more difficult for Charlie to learn, a new procedure was introduced where on training trials initially only one picture was presented and Charlie was asked "Where is OEF (or UI) ?". When performance was stable on this task the original training procedure with both stimuli was reintroduced. During the last training sessions, where two pictures were presented, Charlie responded correctly on 67% of B5-A5 trials and 83% of B6-A6 trials. When tested on A-B (picture-name) relations, Charlie responded correctly on 83% of both A5-B5 and A6-B6 trials. Charlie was 19 months and 8 days when this experiment ended.

Experiment 2 investigated the effect of temporal delays on the arbitrary stimulus relations. Saunders, Saunders and Spradlin (1990) had given a mildly retarded subject reinforced baseline training trials and then retested the baseline and equivalence relations two and three years after the original training. They found that the subject's performance in the absence of feedback was highly accurate, and they suggested the stable performances on these relations were similar to the long term stability of language. In this experiment Lipkens et al wanted to investigate changes or retention of relations across time when training and test performances were not optimal. Charlie had shown mutual entailment in experiment 1 even when performance on the trained relations was not highly accurate. Lipkens et al suggested that performance on the underlying trained relations might improve as Charlie acquired more relational and imitative abilities so Charlie was tested on the set 2 A3-B3, A4-B4 (and vice-versa) relations after two weeks with no additional training. Lipkens et al also repeated experiment 2 with a new set of stimuli, A7-B7 and A8-B8, but this time with delayed testing. This was to investigate whether, if the trained relations were once again weak, the trained and mutual entailment relations would emerge over time.

After the initial training of set 2 A-B (name-picture) and testing of B-A (picture-name) relations with A3, B3, A4 and B4, there was a two week gap before retesting these relations. During this time the B5-A5 and B6-A6 relations in experiment 1 were trained and the A-B relations tested. The set 2 relations (both training and test) were then retested without any additional reinforcement. Charlie now responded 100% correctly on all training and test trials. So even in the absence of reinforcement Charlie's performance had improved dramatically on the test trials and even the training trials where performance had previously been rather weak.

Two new picture-name relations, A7-B7 and A8-B8, were also trained. Charlie responded correctly on 58% of A7-B7 trials and 0% correctly on A8-B8 trials. On the majority of the trials where Charlie's response was scored as incorrect he failed to say anything when asked "What is this?" when shown the sample. The tests for B-A mutual entailment were not given immediately after A-B training. Instead there was a gap of seven days between the last training session and the first test session, when both the training and test relations were probed without additional reinforcement. When tested on the training relations Charlie responded 100% correctly on the A7-B7 relations and 33% correctly on the A8-B8 relations. In the tests for B7-A7 and B8-A8 mutual entailment Charlie responded 100% correctly on all trials. Charlie was 18 months and 11 days old when experiment 2 ended.

Lipkens et al suggested that these data refuted one criticism of operant accounts of language development, that as parents do not continuously reinforce their children's verbal behaviour an operant account is not viable. Lipkens et al believed that the data showed that continuous feedback was not necessary to maintain or improve correct responding on trained or derived relations.

Experiment 3 tested for derived relations of combinatorial entailment. Tests for combinatorial entailment require two sets of related conditional discriminations and several procedures were tried before a successful method for teaching these was found. The relations trained were A-B (picture-name) as in experiments 1 and 2 and A-C (picture-sound). On A-C trials Charlie was shown a picture and asked "What does this say?". After training, tests were given for the emergence of B-C (name-sound) and C-B (sound-name) relations of combinatorial entailment.

The relations trained were the set 1 relations A1-B1 and A2-B2 trained in experiment 1, and two more relations were added to these, A1-C1 and A2-C2. These A-C relations were then tested for the emergence of C-A relations of mutual entailment. The training and testing in this experiment were carried out over a period of 4 months and 13 days. On the tests for mutual entailment Charlie scored 97% correct on B-A (name-picture) relations and 83% correct on C-A (sound-picture) relations. These were overall results with Charlie's performance improving over the time of testing. Charlie showed a similar improvement on the tests for combinatorial entailment. When first tested Charlie scored 0% correct on C-B trials and 62.5% correct on B-C trials. On the final tests Charlie scored 93% correct on C-B trials and 87.5% correct on B-C trials. So when first tested Charlie showed little evidence of derived combinatorial entailment but these relations had emerged

by the end of the testing. However, as Lipkens et al noted, it is not clear if this improvement was due to a developmental trend or merely due to an increase in task familiarity.

Experiment 4 was designed to test if a stimulus exclusion paradigm would generate derived relations. If so, would these derived relations be based on formal/direct properties or on arbitrary relations between the stimuli. Three tests were given, at different stages in the experiment. All the relations tested are described in Fig. 4:2. On the first test two pictures, A9 and A10 were presented. A9 was a picture of a dog and Charlie had already shown that he could name the picture and understood the word "DOG". The other picture, A10, was of a Torosaurus, which Charlie had never seen before and did not know the name for. When the two stimuli were presented Charlie was asked "Where is the DOG ?" or "Where is NEUS?" in random order (name-picture trials). Charlie was then given picture-name trials where one picture was presented and Charlie was asked "What is this?". After this training the same procedure was carried out with another set of stimuli of a known picture, a new picture, a known name and a new name, A11, A12, B11 and B12. Charlie was aged 16 months and 26 days when testing began.

With the first set of stimuli Charlie scored 67% correct on B9-A9 and 100% correct on A9-B9 tests, both known relations. On the excluded B10-A10 (name-picture) relations Charlie scored 100% correct, but 0% correct on the A10-B10 (picture-name) relation. Charlie showed similar performance with the second set of stimuli. On the known relations Charlie scored 83% correct on B11-A11 and 100% correct on A11-B11 relations. On the excluded B12-A12 trials Charlie scored 100% correct but on A12-B12 performance was never correct. Therefore in both sets Charlie clearly demonstrated exclusion in name-picture problems; when given a novel name he pointed to a novel object. However he did not derive a mutual relation between this novel name and the novel object, as his performance was never correct on picture-name trials.

Eight months later (when Charlie was aged 24 months and 13 days) Charlie was given the second test when both sets of the first test were again tested for exclusion. Two new test trials were also given. The two new pictures A10 and A12 were presented and Charlie was asked "Where is NEUS/IEP ?". On this occasion, with the first set of stimuli, Charlie scored 91% correct on B9-A9 and 73% correct on A9-B9, both known relations. On the excluded relations he scored 91% correct on B10-A10 and 18% correct on A10-B10. With the second set of stimuli Charlie

scored 100% correct on both the B11-A11 and A11-B11 known relations. On the excluded relations he scored 90% correct on B12-A12 and 10% correct on A12-B12. On the two new test relations where the novel stimuli A10 and A12 were presented together, Charlie scored 100% on B10-A10 and 90% correct on B12-A12. Lipkens et al suggested that this showed the novel name - novel picture relation in the exclusion trials was not dependent on a familiar figure once the relation had been derived. However this derived relation was not mutual as Charlie's performance was poor on all A-B (picture-name) relations.

For the third test, a third set of stimuli, A13, A14, B13 and B14 were used. This time the B-A (name-picture) and the A-B (picture-name) trials were preceded by some familiar questions. So for instance, before the B-A (name-picture) relations were tested, Charlie might be shown pictures of a cat and a pig and asked "Where is the cat/pig?". He might then be shown stimuli A13 (picture of a baby) and A14 (picture of a chameleon, named a moemoe), and asked "Where is the baby/moemoe?". Before testing the A-B (picture-name) relations, Charlie might be shown a picture of a car and asked "What's this?". He could then be shown A13 (baby) and asked "What's this?", and similarly for A14. Reinforcement was provided on the familiar questions but not the questions involving A13, A14, B13 and B14.

On this task, Charlie showed improvement over time. Using this procedure, on the first set of tests Charlie scored 100% correct on the B13-A13 and A13-B13 known problems. On the excluded relations Charlie scored 83% correct on B14-A14 and 33% correct on the A14-B14 relations. In the last test session, Charlie again scored 100% correct on the known B13-A13 and A13-B13 relations. On the excluded relations Charlie scored 83% correct on B14-A14 and 100% correct on A14-B14 (a substantial improvement on previous A-B (picture-name) relations). Charlie was 27 months and 11 days old at the end of this experiment.

Charlie demonstrated non-verbal exclusion at 17 months old (three months younger than any previous demonstration). Verbal exclusion appeared to emerge over time. In the first test Charlie got none correct, in the second test he averaged 14% correct, by the third test he began at 33% correct and was scoring 100% correct by the end of testing.

The four experiments in this study showed that a variety of derived stimulus relations could be demonstrated with a very young child with a relatively limited

verbal repertoire. Several of these relations were demonstrated at a younger age than in the previously published literature. Lipkens et al suggested that the demonstration of derived stimulus relations in a child as young as 17 months meant that it was unlikely that these relations were dependent on language mediation, as only very simple language processes would have been available to the subject.

The data also suggested that there might be a developmental trend in relational responding. There was a delay of 2 months 9 days between the demonstration of derived name-picture mutual entailment relations and picture-name mutual entailment relations. Similar delays were shown between derived picture-name mutual entailment relations and sound-name/name-sound combinatorial entailment relations. Experiment 3 also showed the gradual emergence of verbal exclusion.

The studies by Devany, Hayes and Nelson (1986) and Barnes, McCullagh and Keenan (1990) had both tried to identify groups of human subjects for whom conditional discrimination training did not seem to result in the emergence of equivalence relations. By comparing the performances of these subjects with performances by subjects who were matched on a number of measures, but who did demonstrate equivalence relations, Devany et al and Barnes et al were able to speculate about the processes underlying equivalence relations. Lipkens, Hayes and Hayes (1993) took a slightly different approach by studying at what stage a young child was able to demonstrate differing aspects of derived relations. In this way they tried to compare the child's performance on tasks of derived equivalence with the cognitive processes available to the child at that time. In this way Lipkens et al were also able to speculate about the possible basis for emergent equivalence relations.

Following research like this, several theories have been put forward to explain these emergent, untaught relations. Each of these theories has been supported or contradicted by varying experimental investigations and these will be considered in greater detail.

Sequence Classes

One area of research which seems to link the paradigms of stimulus equivalence and transitive inference is the investigation into the establishment of sequence classes. As with stimulus equivalence and transitive inference, this research looks at the establishment of new relations as a result of previously established relations.

One of the first investigations into this area was by Lazar (1977). This study came from attempts to expand the usefulness of the matching-to-sample paradigm being developed by Sidman and his co-workers throughout the 1970's. Sidman's experiments had shown that having been directly taught several relations between stimuli via matching-to-sample training, new relations between the stimuli emerged untaught, and the subjects formed appropriate classes of equivalent stimuli on the basis of this training. Lazar's (1977) experiment was an attempt to expand the usefulness of this paradigm to include stimulus relations established in different contexts.

"The matching-to-sample procedure would become of greater importance if it could provide a means by which stimuli gain membership in classes established outside the matching paradigm" (p.382)

Lazar (1977)

The basic plan of this experiment was to establish two stimulus sequence classes, "first" and "second", so that the subjects always responded to the stimuli in one class first and the stimuli in the other class second. The stimuli in these classes were then directly related to new stimuli via a matching-to-sample procedure. Following this training, the new stimuli were tested to see if they had become members in the sequence classes "first" and "second".

Lazar (1977) tested three normally able adult subjects. In the first phase of the experiment the subjects were presented with a pair of visual stimuli, e.g. A1 and A2, and taught to point first to A1 and then to A2 ($A1 \rightarrow A2$), regardless of the spatial position of each stimulus. This training was repeated to establish the sequences $A1 \rightarrow A2$, $B1 \rightarrow B2$, $C1 \rightarrow C2$, $D1 \rightarrow D2$. All three subjects learned these four baseline sequences. The subjects were then presented with mixed pairs, e.g. B1 and A2, to see if this training had established sequence classes, so that the subject's behaviour was controlled by a class of stimuli to which they responded first and by a class to which they responded second. If these classes had been established, given the pair B1 and A2 the subjects would be expected to respond in the order $B1 \rightarrow A2$, given the pair B2 and C1 to respond $C1 \rightarrow B2$ etc. Two subjects (S2 and S3) clearly demonstrated the establishment of the sequence classes "first" and "second". One subject (S1) performed at chance level when initially tested, but following remedial training with different stimuli demonstrated the sequence classes when tested a second time.

Before continuing to the match-to-sample training, the subjects were given sequence pre-training with novel stimuli E1, E2, F1 and F2. This was to check that the subjects consistent sequential responding was a result of the training given rather than a spontaneous tendency to respond sequentially. For all three subjects, performance on the previous baseline and sequence class trials was at or near 100% correct, while performance on the new sequence trials was around chance level. Thus, the subjects were not by themselves producing consistent patterns of sequential responding but rather this responding could be attributed to the training given.

The next phase constituted the main purpose of the experiment. The aim was "to determine whether matching-to-sample would transfer sequence class membership from stimuli in the baseline sequences to the new stimuli" (p.387). The matching-to-sample trials were presented to the subjects in three sets. In all sets the stimuli E1 and E2 or F1 and F2 functioned as the comparison stimuli. In set 1 the sample stimuli were A1, A2, C1, C2, so that A1 and C1 were related to E1 and F1, and A2 and C2 were related to E2 and F2. The subjects also received trials where E1, E2, F1 or F2 functioned as samples as well as comparisons. These trials were to control for possible sequence effects in the matching-to-sample procedure. While the subjects did not have to touch the sample in the matching procedure they did have to observe it before responding to one of the comparison stimuli. This responding to the comparison second might have caused all the new stimuli to become members of the class "second". Giving trials where the new stimuli functioned as both samples and comparisons was designed to prevent this happening. In set 2 of the matching-to-sample training, stimuli B1 and B2 functioned as the sample stimuli and in set 3 the sample stimuli were D1 and D2. The matching-to-sample training was split into these three sets as it was possible that the new stimuli would become members of the critical sequence classes after being related to only a few of the baseline stimuli. Thus, initially the relations in set 1 were trained and then relations from all three sets were tested. If performance on sets 2 and 3 were poor, the set 2 relations were trained and all sets of relations were tested again. If performance on set 3 relations was still poor, the set 3 relations were directly trained.

All three subjects learned the matching to sample relations in set 1 and were tested on the relations for all three sets. All three subjects scored 100% correct on the directly trained relations, e.g. given A1, C1, E1 or F1, they reliably selected E1 or F1, and given A2, C2, E2, or F2 selected E2 or F2. However no subject scored

above chance level on the relations in sets 2 and 3 which had not been trained yet. The BE and BF matching relations in set 2 were then taught. When tested, all subjects responded at 100% on the directly trained set 1 and set 2 relations but subject S1 also responded at 100% on the untrained set 3 relations suggesting that this performance had emerged as a result of the directly trained set 1 and 2 relations. Performance on set 3 was around chance level for subjects S2 and S3 so both these subjects received direct training on the set 3 relations.

The final question in this experiment was whether the matching-to-sample training given would cause transfer of sequence class membership from the original A, B, C and D stimuli used as samples to the new E and F stimuli used as comparisons. Therefore if the subjects were shown the stimuli E1 and E2 would they touch them in the order $E1 \rightarrow E2$. Similarly, would E2, F1 produce the sequence $F1 \rightarrow E2$? Two subjects, S1 and S2 responded at almost 100% on these tests. "Without a direct history of training on these tasks, they responded in predictable sequences, pointing to E1 and F1 first, and E2 and F2 second" (p.389) Subject S3's performance was similar to that on the sequence pre-tests prior to matching-to-sample training, around chance level.

Subject S3 had received the same training that produced transfer with subjects S1 and S2 and had clearly demonstrated the sequence classes "first" and "second" earlier in training. Subject S3 therefore received further testing to see if the matching-to-sample training had established stimulus classes independent of the baseline sequence classes or if the training had failed to establish classes at all. If the sample and comparison had become equivalent through the matching procedure they should have been mutually substitutable as sample and comparison stimuli. Therefore subject S3 was given a series of tests where the stimuli that had previously functioned as samples now functioned as comparisons and vice-versa. If the stimuli in the matching procedure had become equivalent, S3 should have been able to respond appropriately on these tests without further training. When tested, subject S3 responded at 100% correct on the original matching trials but at chance level on the new matching tests, suggesting that the stimuli had not become equivalent and were not interchangeable. This made the failure of sequence class membership to transfer to the new E and F stimuli more understandable.

This experiment showed that the matching-to-sample/equivalence paradigm being developed by Sidman could be a useful way of establishing class membership. "Matching-to-sample can extend class membership to new stimuli even when the

stimulus classes are originally established outside the matching context" (p.391). However Lazar did note that this procedure did not produce perfect transfer of class membership as only two out of three subjects eventually displayed this transfer. Also, there were good theoretical grounds for expecting the subjects to display transfer of function to all stimuli after receiving only set 1 matching training, yet training on at least set 1 and set 2 matching trials was necessary before any transfer was shown. However, despite this, the matching paradigm had proved to have some utility in extending class membership and Lazar suggested that there were good reasons for continuing investigation of this type. For instance, it may be relevant to the emergence of aspects of language; "a behavioral repertoire, much of which seems to emerge without direct training" (p.391). Lazar suggested that possibly words that occupy equivalent positions could be members of the same grammatical class; "this might, for example, permit a child to combine a separately learned adjective and noun appropriately, even though that particular combination has never been directly trained" (p.392). Lazar was not suggesting that the procedure detailed in this experiment explained or described the development of linguistic behaviour, but rather that this procedure might provide a way to examine some of the pre-requisite skills.

Lazar and Kotlarchyk (1986)

A study by Lazar and Kotlarchyk (1986) was designed to examine if normally able pre-school children would show similar behaviour to the adults in the study by Lazar (1977) and also extended the paradigm used by Lazar (1977) by bringing the sequence classes under second order control.

"The objective . . . was to ascertain whether stimulus-class formation could serve as a basis for new sequence behavior in pre-school children, and in addition show that sequences can be brought under second-order stimulus control" (p.206)

In this experiment the subjects were four normally able children (1 male, 3 female) aged from 5:10 to 6:4. In the first phase of the experiment, the subjects were taught to match Greek letter comparisons to colour samples. For example, when the sample was red, the subjects were taught to select stimulus A1, and when the sample was green, to select stimulus B1. When the sample was red the subject was to select A2 and when green to select B2 etc. All subjects were able to learn the baseline matching tasks. The purpose of this training was to establish two 5-member stimulus classes: red, A1, A2, A3, A4 and green, B1, B2, B3, B4. The subjects then received tests to see if these classes of equivalent stimuli had formed.

If the stimuli had formed equivalence classes it should have been possible for the subjects to match the stimuli within each class to each other without further training, e.g. given stimulus A1 they should select stimulus A2 rather than B2 etc. The subjects were therefore expected to match all the A stimuli to each other and all the B stimuli to each other. When tested on this task the subjects all performed at nearly 100% correct on the baseline matching tasks and demonstrated the formation of classes of equivalent stimuli in between 1 and 4 sessions.

Having shown that the matching to sample training had established classes of equivalent stimuli, the subjects then received sequence training with the red and green stimuli that had previously functioned as samples. The sequence training was under the control of two auditory stimuli, so that in the presence of tone 1 the subjects had to touch the stimuli in the order red→green, and in the presence of tone 2, touch the stimuli in the order green→red regardless of the spatial position of the two stimuli. All four subjects learned these second-order sequences. Having learned these sequences "the experimental question was whether the ordinal-class properties of red and green would be acquired by corresponding members in the two equivalence classes" (p.212). In other words, if the A stimuli had acquired the function of the red stimulus and the B stimuli had acquired the function of the green stimulus, then in the presence of tone 1, given a pair of stimuli such as A1 and B1, the subjects should touch the stimuli in the order A1→B1, and in the presence of tone 2, touch the stimuli in the order B1→A1. This should occur whatever the combination of A and B stimuli, e.g. A2 and B2, A1 and B4 etc. The subjects were therefore given tests for the emergence of these ordinal properties. For all four subjects baseline performances were nearly perfect and performances on the untrained, second order-sequences were almost 100% correct within 2 sessions. Therefore the order functions of the green and red stimuli had transferred to the other members of their respective stimulus classes without an explicit training history.

This experiment effectively replicated Lazar's (1977) findings with adults with pre-school children and also extended Lazar's (1977) procedure by bringing the sequence classes under contextual control. As Lazar stated, these findings provided support "for the notion that stimuli which have become equivalent in the context of matching to sample can also share membership in classes established outside the matching paradigm" (p.213).

Lazar and Kotlarchyk stated that a reason for interest in the characteristics of equivalence classes was "The resemblance between simple linguistic performances and the outcomes of arbitrary matching to sample" (p.213). The results from Sidman (1971) and Sidman and Cresson (1973) showed that pictures and words had become mutually substitutable as a result of the training given and Lazar and Kotlarchyk (1986) suggested that this meant the stimuli had acquired "semantic" properties. Similarly the results from Lazar (1977) and Lazar and Kotlarchyk (1986) showed that matching to sample extended membership in classes with "structural" properties - e.g. sequential relations. Lazar and Kotlarchyk (1986) felt that this represented an approximation of grammar by showing that stimuli could share multiple membership in ordinal classes in the same way that a word "can function in more than one position in a sentence depending on the context" (p.213). Lazar and Kotlarchyk pointed out however that they were not trying to teach a form of grammar but rather trying "to establish procedures that examine generative sequence behaviour" (p.214).

Wulfert and Hayes (1988)

Another study by Wulfert and Hayes (1988) examined transfer of sequential function through equivalence classes. As with Lazar (1977) and Lazar and Kotlarchyk (1986), Wulfert and Hayes were interested in how performances on equivalence classes and sequential classes may be related to the development of semantics and syntax. Wulfert and Hayes therefore carried out two experiments, with different training histories but designed to produce functionally identical responses. Experiment 1 examined the transfer of sequential responding through equivalence classes (phase 1) then through conditional equivalence classes (phase 2), then the transfer of conditional sequential responding through conditional equivalence classes (phase 3). Experiment 2 established equivalence classes and sequential responding as in experiment 1 (phase 1). Then the order of training was reversed compared to that in experiment 1, with the sequence training being brought under conditional control (phase 2) and then the equivalence classes being brought under conditional control (phase 3). Wulfert and Hayes argued that it was unlikely that verbal behaviour was acquired in a strict sequential fashion. Therefore if the subjects' performances on these tasks varied according to the order of training given, then it was unlikely that these experiments were investigating the processes operating in syntax and semantics. However, "a demonstration that functionally identical complex response patterns can emerge from different histories would increase the plausibility of the proposed account" (p.131) i.e. that

these experiments were investigating some of the processes operating in syntactic and semantic behaviour.

Eight college students took part in this study, four in experiment 1 and 4 in experiment 2. All four subjects in both experiments showed the transfer of the sequence response, whatever the order of training, although in both experiments there was considerable intersubject variability in the amount of training and testing needed to complete the experiment.

Wulfert and Hayes suggested that if equivalence relations are semantic in nature, as proposed by Sidman (1986), then the performances shown in these experiments may parallel the emergence of simple two-word utterances in a young child. If the child has learned several concepts (e.g. colours, toys food) he or she may then be able to combine these in a number of ways, from only a few directly trained instances.

"If the assumption is correct that verbal concepts originate from the participation of verbal stimuli in equivalence or other relational classes, the emergence of new, untrained combinations of elements of these classes seems to lose its mysterious quality" (p.139)

Wulfert and Hayes did note that their subjects were adults who already had well developed language skills and sophisticated behavioural repertoires and this may have facilitated their performance on these tasks, so their performance may be significantly different to that of a young child who is still developing semantic and syntactic skills. However they felt this research demonstrated that "the environment can establish rather sophisticated control over sequential responses to symbols and that from few trained instances a very large and flexible number of untrained sequences can arise" (p.140) and that continued research of this type would help to elucidate some of the processes involved in verbal behaviour.

These studies by Lazar (1977), Lazar and Kotlarchyk (1986) and Wulfert and Hayes (1988) were very much rooted in the matching to sample/stimulus equivalence paradigm that had been developed by Sidman and his co-workers. Lazar (1977) and Lazar and Kotlarchyk (1986) in particular were not so much interested in the features of sequence classes themselves but more in sequence classes as away of examining transfer of function in an equivalence paradigm. Other researchers however, have made a more detailed study of the development of sequence classes and the emergent properties within these classes. In many ways,

this research had drawn more on findings from transitive inference research than from stimulus equivalence.

Robert Stromer and Harry Mackay carried out much of this research into the development of sequence classes. Their work aimed to “extend the analysis of sequential performances further and to assess the defining properties of ordinal stimulus classes” (Stromer and Mackay, 1990, p.2). Sidman and Tailby’s (1982) definition of stimulus equivalence requires that it should be possible to demonstrate the properties of reflexivity, symmetry and transitivity if true equivalence relations have formed. Sequence relations or ordinal relations also demonstrate transitivity but in this case are asymmetric. Perhaps the best example of this is in transitive inference. If the subject is taught $A > B$ (A “greater than” B) and $B > C$, then through transitivity the subject can infer the relation $A > C$, but this requires that the relation is asymmetric, A cannot be “greater than” ($>$) and “less than” ($<$) C at the same time. In the case of sequential relations this asymmetry is described by the terms “before” and “after” as the subjects are taught to touch, or respond to, the stimuli in a specific order.

Stromer and Mackay (1990)

Stromer and Mackay (1990) carried out a study investigating some of the characteristics of sequence relations. Bryant and Trabasso (1971) and Lutkus and Trabasso (1974) had established a sequence of five stimuli $A > B > C > D > E$ by training four overlapping two-term sequences $A > B$, $B > C$, $C > D$, $D > E$. Other two-term sequences were derived from this. Stromer and Mackay asked if directly training a five term series would produce the two-term sequences found by Bryant and Trabasso (1971) and Lutkus and Trabasso (1974). A normally able adult was taught to touch five arbitrary visual forms in the order $A1 \rightarrow A2 \rightarrow A3 \rightarrow A4 \rightarrow A5$ and subsequently produced ten two-term subsequences, $A1 \rightarrow A2$, $A1 \rightarrow A3$, $A1 \rightarrow A4$, $A1 \rightarrow A5$, $A2 \rightarrow A3$, $A2 \rightarrow A4$, $A2 \rightarrow A5$, $A3 \rightarrow A4$, $A3 \rightarrow A5$, $A4 \rightarrow A5$. Stromer and Mackay suggested this showed that “the stimuli were related to one another by virtue of their relative positions in the five term sequence trained originally” (p.3).

If responding to the stimuli was controlled by serial position then it was possible that stimuli occupying the same serial position in independent sequences would form a class. To test this the subject was taught a second five-term sequence, $B1 \rightarrow B2 \rightarrow B3 \rightarrow B4 \rightarrow B5$ and again demonstrated the two-term subsequences. The subject was then given 20 probe trials of two-term subsequences mixed from the five-term sequences, e.g. $A1 \rightarrow B2$, $B2 \rightarrow A3$ etc. Performance on these

subsequences was consistent with the experimentally designated sequences, suggesting that the stimuli had formed classes based on their position in the sequence and were therefore mutually substitutable. The emergence of the mixed two-term subsequences $A2 \rightarrow B4$ and $B2 \rightarrow A4$ suggested that relations in the sequences were transitive. As in the transitive inference experiments the relation between the stimuli in the second and fourth positions in the sequence was the best indicator of transitivity as it did not involve the endpoints of the series and so did not involve absolute responding. However even the demonstration of $A2 \rightarrow A4$ and $B2 \rightarrow B4$ was not reliable evidence of transitivity as the subject had direct visual exposure to the stimuli involved through the five term sequences. However in the mixed $A2 \rightarrow B4$ and $B2 \rightarrow A4$ subsequences the stimuli had never appeared together before and so provided good evidence for transitivity.

A third experimental phase extended these findings by establishing a new "second" (X2) and a new "fourth" (X4) stimulus. These were trained by teaching the sequences $A1 \rightarrow X2$, $X2 \rightarrow A3$, $B3 \rightarrow X4$ and $X4 \rightarrow B5$. Performance was assessed on a number of two-term subsequences, both adjacent e.g. $X2 \rightarrow B3$ and non-adjacent $X2 \rightarrow A4$ etc, and on four five-term sequences e.g. $A1 \rightarrow X2 \rightarrow A3 \rightarrow A4 \rightarrow A5$. Performance on all these was very good and improved with repeated testing. Another set of "second" and "fourth" stimuli was established (Y2 and Y4) and once again probe performances were almost perfect. This included mixed two and five-term probes of the new stimuli (e.g. $X2 \rightarrow Y4$ and $A1 \rightarrow Y2 \rightarrow A3 \rightarrow X4 \rightarrow A5$), giving further support to the proposition that the stimuli formed sequence classes and that the five-term sequences were transitive. In the final phase of the experiment Stromer and Mackay tested whether training a series of adjacent two-term sequences would produce probe performances indicating transitivity and the formation of ordinal relations. The five-term A and B sequence baselines were maintained and the subject was taught $Z1 \rightarrow Z2$, $Z2 \rightarrow Z3$, $Z3 \rightarrow Z4$, $Z4 \rightarrow Z5$. Subsequent probes were consistent with the establishment of ordinal relations - both two-term, $Z2 \rightarrow Z4$ and five-term probes, $Z1 \rightarrow Z2 \rightarrow Z3 \rightarrow Z4 \rightarrow Z5$, and mixed probes $A2 \rightarrow Z4$ etc.

This study examined some of the conditions under which sequence classes will develop and demonstrated that the stimuli in these classes are mutually substitutable. This study extended the procedure used by Lazar (1977) who trained $A1 \rightarrow A2$, $B1 \rightarrow B2$, $C1 \rightarrow C2$ and $D1 \rightarrow D2$ and then demonstrated mixed sequences such as $B1 \rightarrow C2$, $A1 \rightarrow D2$. By using aspects of the procedures used to investigate

transitive inference, Stromer and Mackay were able to make a more detailed analysis of the development of sequence classes.

Stromer and Mackay (1992)

Stromer and Mackay developed this analysis of sequence classes with another study in 1992 which examined conditional stimulus control of trained sequence classes. In this experiment the subjects were six normally able children, aged 8 to 9 years. The children were taught to produce two five-term sequences by teaching them to select abstract visual stimuli in a specific order. The stimuli were presented on a computer screen, with all the stimuli for each sequence presented mixed together in a choice area. Selecting a stimulus moved it to a sequence construction area. The children were taught the series $A1 \rightarrow A2 \rightarrow A3 \rightarrow A4 \rightarrow A5$ and $B1 \rightarrow B2 \rightarrow B3 \rightarrow B4 \rightarrow B5$. When both these series had been established, the A sequence was brought under conditional control. When the printed word BIF was displayed at the top of the computer screen, the subjects were reinforced for touching the stimuli in the order $A1 \rightarrow A2 \rightarrow A3 \rightarrow A4 \rightarrow A5$. When the printed word NUK was displayed the subjects were reinforced for touching the stimuli in the reverse order $A5 \rightarrow A4 \rightarrow A3 \rightarrow A2 \rightarrow A1$. The subjects then received probe trials to see if the conditional control by the printed words BIF and NUK had transferred to the B sequence stimuli. Four subjects immediately produced the appropriate B sequences under conditional control and a fifth subject (LA) showed improved accuracy when the test session was repeated. One subject (LD) always touched the stimuli in the order trained originally. The subjects who had demonstrated appropriate conditional control then received multiple-substitution probes - two or three stimuli that occupied the second, third and fourth positions in the A sequences replaced the comparable stimuli in the B sequences, and vice versa. Four of the subjects immediately produced these mixed sequences accurately. The fifth subject (LA), who had required a second test session with the B stimuli, was always incorrect on the mixed sequences.

The two subjects, LA and LD, who did not initially succeed on the test probes then received further training and testing to see if this would improve their performance. Subject LA had shown conditional control on the B sequence but had failed to respond correctly on the multiple-substitution probes. LA was given single substitution probes where one of the B stimuli substituted for one of the stimuli in the A sequence and vice versa - e.g. $A1 \rightarrow B2 \rightarrow A3 \rightarrow A4 \rightarrow A5$. Subject LA was still below criterion on this task so these single substitution sequences were trained

directly. Following this training LA was given the multiple substitution probes again and this time scored 81% correct.

Subject LD had failed to show transfer of conditional control from the A to the B sequences, so this conditional sequencing was taught directly and then the subject was tested on the multiple substitution probes. Performance on these probes was well below criterion, so as with subject LA, LD was taught the single substitution sequences and the multiple substitution probes were tested again. At best LD achieved only 31% correct.

This experiment extended the work on sequence classes by Lazar and Kotlarchyk (1986) and Wulfert and Hayes (1988) by "demonstrating transfer of conditional control without first using matching to sample tasks to establish equivalence relations between the stimuli involved" (p.910) As in the previous experiment (Stromer and Mackay, 1990), this experiment showed that stimuli occupying the same serial position in separate sequences were mutually substitutable, suggesting that the training had established functional stimulus classes based on the arbitrarily assigned positions of the stimuli.

However, subject LA's data suggested that the establishment of sequence classes and the transfer of conditional control may be independent of each other. This subject demonstrated transfer of conditional control but required additional training before performing correctly on the multiple substitution probes.

Stromer and Mackay (1993)

A further study by Stromer and Mackay (1993) repeated some of the procedures already reported (Stromer and Mackay, 1990) under more rigorous conditions, and with a larger number of subjects. Experiment 1 tested whether subjects taught to touch five stimuli in a particular order would also touch pairs of these stimuli in an order consistent with training. If the subjects were able to respond in this way it would suggest that the training had established ordered relations among the stimuli in the sequence rather than a simple discriminative chain. This experiment also tested whether stimuli occupying the same position in different five term sequences would prove to be mutually substitutable, suggesting that the training had established functional stimulus classes. The emergence of these mixed sequences would also provide further evidence against a chaining account of the procedure.

In experiment 2 of this study, an overlapping two-term sequence training procedure, similar to the procedures used in experiments on transitive inference, was used. The results from the transitive inference experiments (e.g. Bryant and Trabasso 1971, Gillan 1981, Lutkus and Trabasso 1974, McGonigle and Chalmers 1977, Fersen, Wynne, Delius and Staddon 1991) had suggested that this procedure also determines the position of each stimulus in the overall series. To test this, after training with four two-term adjacent sequences, the subjects were tested for the emergence of non-adjacent sequences, and also for the emergence of a new five-term sequence. Probes were also given to see if mixed sequences would emerge consisting of stimuli trained by a five-term training procedure and stimuli trained using the overlapping two-term procedure. This would provide further evidence for the development of sequence classes, and against a chaining account of the sequence relations.

The subjects in these experiments were two normally able children aged 9 and 10 years, and five normally able adults in their mid-twenties. In experiment 1, two five-term stimulus sequences, $A1 \rightarrow A2 \rightarrow A3 \rightarrow A4 \rightarrow A5$ and $B1 \rightarrow B2 \rightarrow B3 \rightarrow B4 \rightarrow B5$ were trained and the subjects were tested to see if they produced appropriate two-term subsequences, and also if they would produce appropriate mixed A/B two-term and five-term sequences. Two adults (MM and AM) and one child (AS) participated,

Subject MM received training on the A and B five-term sequences. MM then produced all the two-term A sequences and all the two-term B sequences correctly. MM also responded correctly on all the mixed two-term A/B sequences - even though these probes involved mixtures of stimuli that had not occurred together in training.

Subject AS was also taught the five-term A and B sequences but this subject was given the two-term mixed A/B probes first and the two-term A and B probe sequences next. Performance on the two-term A/B probes was correct for sequences beginning with an A stimulus but was not correct for sequences beginning with a B stimulus. However performance on the two-term A and B sequences was perfect. When given five-term mixed A/B single substitution probes AS's performance was consistent with training only on probes where A5 and B5 were substituted for each other. Because of this inconsistent performance these mixed A/B five-term single substitution sequences were trained directly. Following this training, performance on the five-term multiple substitution sequences was

nearly perfect. Before retesting the two-term mixed A/B sequences the eight adjacent two-term A and B sequences were directly trained and added to the subjects baseline to see if this would facilitate correct responding. Performance on the five-term mixed A/B multiple substitution sequences and on the two-term mixed A/B sequences was then reassessed. As before, performance on the two-term sequences was correct only for those sequences beginning with an A stimulus. Performance was nearly perfect on the five term multiple substitution sequences.

Subject AM was taught the five-term A and B sequences. This subject was then tested on the five-term multiple substitution A/B sequences, then on the two-term A/B sequences, then on the five-term multiple substitution sequences once more. Performance on all these two and five term A/B probe sequences was nearly perfect.

In summary then, following training on two five-term A and B sequences, subject MM was able to respond correctly on mixed A/B two and five term probes, suggesting the formation of classes of sequence stimuli. Subject AM's performance replicated and extended these results. AM also produced appropriate mixed two and five term sequences, but as AM had not received testing on the within-sequence two-term A and B sequences, it would seem that these within-sequence performances were not necessary for occurrence of the mixed performances.

Both subject MM and AS had demonstrated appropriate two-term A and B sequence performance following the training of independent five-term A and B sequences. However, this training appeared to have established sequence classes for subject MM but not for subject AS. Initially AS's performance was incorrect for both five-term and two-term A/B sequences. Training the five-term single substitution sequences resulted in accurate performance on the five-term multiple substitution sequences, but this effect was limited to the five-term sequences. Performance on the two-term A/B sequences remained incorrect, even though performance on the two-term A and B sequences was accurate. This suggested that subject AS could discriminate the A and B stimuli but did not treat the stimuli occupying the same position in separate sequences as mutually substitutable.

The fact that subjects could respond correctly on mixed probes and within sequence probes suggested that the relations between stimuli established during training involved more than simply chaining. However, as Stromer and Mackay pointed out

"The difference between subjects' performances on mixed probe trials emphasizes that the contingencies involved in the independent training of 5-term sequences does not *require* that stimuli that occupy the same position in different sequences become mutually substitutable" (p.118, italics in original).

In experiment 2, subjects were directly taught one five-term sequence (A or B, e.g. $A1 \rightarrow A2 \rightarrow A3 \rightarrow A4 \rightarrow A5$) and the four adjacent two-term sequences that were its components ($A1 \rightarrow A2$, $A2 \rightarrow A3$ etc.). They were also taught four adjacent two-term sequences using five new C stimuli; $C1 \rightarrow C2$, $C2 \rightarrow C3$, $C3 \rightarrow C4$, $C4 \rightarrow C5$. The subjects then received probes to assess performance on: the five-term C sequence that could be derived from the two-term relations trained, the possible five term multiple substitution sequences, and for some subjects, two-term sequences involving the C stimuli. The subjects were three adults (SM, MK, DE) and one child (TJ).

Subject TJ was taught the five-term B sequence and its components and the four adjacent two-term C sequences. TJ immediately produced the five-term C sequence when probed. Performance was also assessed on five-term multiple substitution B/C sequences. This performance was almost perfect.

Subject SM was taught the five-term A sequence and its component parts and the adjacent two-term C sequence. SM was perfect on the five-term C sequence probes. SM was then given probes on five-term multiple substitution A/C sequences and on the two-term sequences $A2 \rightarrow C4$, $C2 \rightarrow A4$, $A2 \rightarrow A4$, $C2 \rightarrow C4$. Performance was perfect on all these probes.

Subject MK was taught the five-term A sequence and its components and the adjacent two-term C sequences. MK produced the correct five-term C sequence on only 2/8 trials. This five-term C sequence was added to the baseline sequences and MK was tested on the five-term multiple substitution A/C sequences. On these sequences MK always touched the A stimuli, in the correct order, first, and then the C stimuli in the correct order. This subject then received training with adjacent two-term A/C stimuli to try and produce the mixed probe sequences; $A1 \rightarrow C2$, $C2 \rightarrow A3$, $A3 \rightarrow C4$, $C4 \rightarrow A5$, $C1 \rightarrow A2$, $A2 \rightarrow C3$, $C3 \rightarrow A4$, $A4 \rightarrow C5$. The subject was then re-tested on the mixed five-term A/C probes and tested on the two-term sequences $A2 \rightarrow C4$, $C2 \rightarrow A4$, $A2 \rightarrow A4$, $C2 \rightarrow C4$. Performance on the five-term and two-term probes was perfect.

Subject ED was taught the five-term A sequence and its components and the two-term adjacent C sequences. Initially none of the five term C probes were correct as the subject always placed C1 at the end of the sequence rather than the beginning. After a three day break, the baseline relations were retrained and performance was perfect when tested again. When the five-term A/C sequences were tested ED showed the same pattern of responding as MK; selecting the A stimuli in order first and then the C stimuli in order. ED was then taught a series of adjacent two-term sequences with new D stimuli in the same way as the C stimuli had been taught. When tested on the five-term D sequence ED's performance was consistent with training. ED was then tested on mixed five-term A/D sequences but showed the same pattern of responding as with the A/C stimuli, touching the A stimuli first and then the S stimuli.

As with subject MK, ED was taught adjacent pairs of two-term A/C sequences to try and produce mixed sequences. ED was then tested on five-term A/C and A/D probes and on eight mixed two-term probes: A2→C4, A2→D4, C2→A4, C2→D4, D2→A4, D2→C4, C2→C4, D2→D4. Performance on the mixed five-term probes was highly accurate and perfect on the two-term probes.

In summary then, following training with adjacent two-term C sequences, two subjects (TJ and SM) immediately derived the five-term C sequences, suggesting that the relations established among the stimuli were transitive. For these two subjects the C stimuli also proved to be interchangeable with the A or B stimuli in the sequence trained previously, suggesting that the training had established sequence classes.

For the other two subjects (MK and ED), the new five-term C sequence emerged only after repeating the probes or following additional training with new stimuli. These subjects also required training with adjacent two-term A/C sequences before they were able to produce mixed five-term A/C sequences. As with subject AS in experiment 1, the subjects' initial performances on the mixed five-term probes demonstrated that the training/test contingencies in this experiment do not *require* the production of mixed sequences of stimuli. Both subjects selected the A stimuli first, in order, and then the C stimuli in order (e.g. A1→A3→A5→C2→C4 or A2→A4→C1→C3→C5). These performances were consistent with the relations established by training, and showed that the subjects discriminated the sets of stimuli, but also showed that at this initial test the training procedures had not caused the formation of sequence classes of interchangeable stimuli.

The results from this experiment replicated and extended some of Stromer and Mackay's preliminary findings (Stromer and Mackay, 1990). The more rigorous testing procedure and greater number of subjects allowed for a more detailed analysis of the results and consideration of the conditions underlying the formation of sequence classes. The two experiments demonstrated that:

1. The training that directly established each five-term sequence may also enable accurate performance on two-term probe sequences involving the same stimuli - a finding which is consistent with some results from non-human subjects (D'Amato and Colombo, 1988).
2. Establishing two separate five-term sequences may cause the emergence of stimulus classes consisting of stimuli that occupied the same position in the different sequences - extending prior findings (Lazar 1977, Lazar and Kotlarchyk 1986, Wulfert and Hayes 1988).
- 3 Following training on adjacent two-term sequences, subjects may demonstrate the emergence of a new five-term sequence, and the stimuli in this sequence may also become members of sequence classes - results which complement and extend findings from transitive inference studies (Bryant and Trabasso 1971, Lutkus and Trabasso 1974).

Stromer and Mackay went on to consider what conditions might underlie or explain performances on these tasks.

The training procedure in experiment 1 could have established simple behaviour chains. However Stromer and Mackay pointed out that it does not seem possible that chaining provides a sufficient explanation for the performances demonstrated. For example, on many of the two term probe trials (e.g. $A_2 \rightarrow A_3$, $A_2 \rightarrow A_4$) the choice pool did not contain the initial stimulus in these sequences (A_1), and the training could not have established A_2 , A_3 etc. as a discriminative stimulus. Also, many of these probe trials demonstrated accurate performance on non-adjacent pairs (e.g. $A_2 \rightarrow A_4$) for which chaining does not provide a sufficient explanation for accurate performance. Probably the most conclusive evidence against a chaining account of these performances is the emergence of the mixed chains. "Because the stimuli for the different A and B sequences never appeared together in training, the contingencies that could have established mixed chains never occurred" (p.126)

Another explanation for the performances might have been based on the simple and conditional discrimination functions that may have been served by the stimuli in each baseline sequence. Performance on the two-term sequences involving stimuli A1 and/or A5 might be explained by the simple discrimination function of these stimuli. In training, when other stimuli from the A sequence were present, selection of A1 was always reinforced and selection of A5 was never reinforced. In that case it would not be surprising if A1 was always selected first on probes trials and A5 always selected last. A similar history would apply to stimuli B1 and B5 in probes with B sequence stimuli and could explain performance on A/B probes containing stimuli A1, A5, B1 and B5.

Another possible explanation might be that the sequence training actually established a set of conditional discriminations involving the stimuli in that sequence. For example, having just selected A1 select A2, having just selected A2 select A3 etc. This would be sufficient to explain accurate performance on the two-term sequences A1→A2, A2→A3, A3→A4, A4→A5. However this account cannot explain accurate performance on the non-adjacent two-term sequences (e.g. A2→A4) and the subjects' success on these probes would suggest that training must have established more than these conditional relations. Once again, the emergence of accurate performance on the A/B probe trials suggests that the establishment of conditional relations is insufficient to explain the results obtained.

"Because the A and B stimuli never appeared together during training, the contingencies that could have established the specific conditional discriminations that would be required for performance on mixed probes never occurred " (p.127)

Stromer and Mackay suggested that research on transitive inference might provide an explanation for one aspect of the derived performances. The training procedure may have established transitive relations between the stimuli in the sequence which would allow successful performance on, at least, the within sequence probes. For instance, as in transitive inference tasks, correct production of the two-term sequence A2→A4 might be based on transitivity of the conditional relations A2→A3 and A3→A4 (therefore A2→A4). The best evidence for the establishment of these relations comes from the procedure used in experiment 2. In this experiment the procedure established the five-term C sequence by training adjacent pairs of C stimuli. The emergence of the five-term C sequence and the correct performance on the non-adjacent two-term probes provided evidence that the training had established transitive relations between the stimuli.

The other feature which Stromer and Mackay suggested might help explain the results obtained was the development of sequence classes. The best evidence for the development of these classes comes from the emergence of new mixed sequences of stimuli. The subjects were able to touch the stimuli presented, in an order consistent with training, even though these stimuli had never appeared together before.

"Whether training used the 5-term or the adjacent-pairs procedure, the stimulus control that it often engendered was not restricted to the specific set of stimuli involved. . . . Apparently, the stimuli that occupied the same ordinal position in different sequences came to form classes of functionally equivalent stimuli."
(p.128)

Not all subjects in either experiment immediately responded correctly on the basis of the sequence training given. However, for those subjects who did respond correctly on probe trials following the training of one or two stimulus sequences, the establishment of transitive relations and sequence classes may explain the emergence of new performances, not required as a result of the contingencies arranged in training.

Stromer, Mackay, Cohen and Stoddard (1993)

A study by Stromer, Mackay, Cohen and Stoddard (1993) replicated the procedures of Stromer and Mackay's 1990 and 1993 experiments using adults with learning disabilities as subjects, to see if these subjects would demonstrate similar emergent performances. A study by McManis (1969) had suggested that it might not be possible to establish ordinal relations in subjects with a mental age less than 10 or 11 years. However studies of transitive inference by Bryant and Trabasso (1971) and Lutkus and Trabasso (1974) suggested that ordinal relations could be established in subjects with mental ages of 5 to 6 years. Stromer et al used the non-verbal sequence production tasks developed in Stromer and Mackay's previous experiments to investigate ordinal learning in subjects with low mental ages. "The goal was to assess with these individuals whether sequence-production training would establish performances based on the sequential relations among stimuli" (p.244).

The subjects in this experiment were two adults with learning disabilities (BF and AP) and one normally able child (DC). BF (aged 45) had Down's Syndrome. Her Peabody Picture Vocabulary Test (PPVT) age-equivalent score was 2-10 (years-

months). AP (aged 61) had a PPVT age-equivalent score of 7-7. The pre-school child (DC) was 5 years old. The stimuli used were abstract visual shapes, presented by a computer with a touch-sensitive screen.

All three subjects were taught the five-term A sequence $A1 \rightarrow A2 \rightarrow A3 \rightarrow A4 \rightarrow A5$ and were then tested for performance on two-term A sequence probes (e.g. $A1 \rightarrow A3$, $A2 \rightarrow A4$). They were then taught the five term B sequence $B1 \rightarrow B2 \rightarrow B3 \rightarrow B4 \rightarrow B5$ and tested on the two-term B sequence probes. The subjects were also tested on mixed two-term A/B probes. The two subjects with learning disabilities (BF and AP) were also given probes to assess performance on single and multiple substitutions on five-term mixed A/B sequences.

Subject BF was taught the five-term A sequence and given the two-term A sequence probes. BF's performance was good on probes that required her to touch A1 first (e.g. $A1 \rightarrow A3$, $A1 \rightarrow A4$) and on non-adjacent probes (e.g. $A2 \rightarrow A4$, $A3 \rightarrow A5$). However on the adjacent stimulus probes (e.g. $A2 \rightarrow A3$, $A3 \rightarrow A4$) her responding was about chance level. BF was then taught the five-term B sequences and given the two-term B sequence probes. BF also received three-term probes (e.g. $B1 \rightarrow B3 \rightarrow B4$, $B2 \rightarrow B3 \rightarrow B5$) to see if additional contextual support might produce higher probe scores. Performance on the two-term B sequences was very similar to those on the two-term A sequences. Performance was perfect for probes beginning with B1 and generally higher on non-adjacent probe pairs than on adjacent probe pairs. On the three term probes BF's performance was highly accurate on probes beginning with B1, but on probes where B2 or B3 had to be touched first performance was often not consistent with training.

In an attempt to improve BF's performance differential reinforcement was provided on the two-term probes. This resulted in a considerable improvement in accuracy but errors continued to occur on the $2 \rightarrow 3$ and $3 \rightarrow 4$ probes so BF was directly taught the two term sequences $A2 \rightarrow A3$, $A3 \rightarrow A4$, $B2 \rightarrow B3$, $B3 \rightarrow B4$. BF was then presented with two-term mixed A/B probes (e.g. $A2 \rightarrow B3$, $B1 \rightarrow A4$). There were twenty possible A/B two term combinations and in the first probe session BF responded correctly on 19 out of 20 probes ($B4 \rightarrow A5$ incorrect). Immediately after the two-term mixed A/B probes BF was given five-term mixed A/B probes. These were both single substitution probes, where each of the A stimuli replaced its corresponding B stimulus in the sequence (e.g. $A1 \rightarrow B2 \rightarrow A3 \rightarrow A4 \rightarrow A5$) and vice versa, or multiple substitution probes, where two or three of the stimuli assigned to the second, third or fourth positions of the original sequences were substituted for

one another (e.g. $A2 \rightarrow B2 \rightarrow B3 \rightarrow A4 \rightarrow A5$). On the single substitution probes BF's performance was perfectly consistent with training and on the multiple substitution probes responses on 6/8 trials were consistent with training.

Subject AP was taught the five-term A sequence and given the two-term A sequence probes. Performance was highly accurate on 9/10 two-term probes ($A3 \rightarrow A5$ performance below 0.83). AP was then taught the five-term B sequence and was tested on both the A and B two-term sequences. Performance was perfect on the two-term A sequences and correct on 8/10 two-term B sequence probes. AP was then tested on the two-term mixed A/B probes. Performance was perfect for all $A \rightarrow B$ probe types and correct on 8/10 $B \rightarrow A$ probe types. Finally AP was tested on single and multiple substitution five-term mixed A/B probes. AP's performance was almost perfectly consistent with training, responding correctly on 9/10 single substitution probes and all multiple substitution probes.

Subject DC was taught the five-term A sequence and subsequently responded correctly on all two-term A sequence probes. DC was then taught the five-term B sequences. When tested on the two-term B sequence probes DC's performance was consistently better when the probes contained either of the end stimuli B1 or B5. This performance may have been due to DC noting that the probes (all two-term) were never reinforced. To prevent this problem in further probe trials the adjacent two-term A and B sequences were added to DC's reinforced baseline trials. This was designed to make the two-term probes less discriminable with respect to not producing reinforcement. DC was then tested on the two-term mixed A/B probes. Performance on these probes was virtually always consistent with prior training.

In summary then, two adults with learning disabilities and one normally able pre-school child were taught two five-term sequences. Following this training, all three subjects were able to produce novel two-term mixed sequences (e.g. $A2 \rightarrow B4$) that were consistent with the serial position of the stimuli in their baseline training. One adult (AP) and the child (DC) were usually capable of producing most of the two-term sequences (e.g. $A2 \rightarrow A3$, $B2 \rightarrow B4$) that made up the baseline five-term sequences. The other adult (BF) responded appropriately on many of these two-term sequences but made repeated errors on others, especially those involving stimuli occupying the second and third positions in the five-term baseline). However once these relations had been established (by direct training if necessary) all three subjects immediately produced almost perfect mixed two-term probe performances.

As with previous research (Stromer and Mackay, 1990, 1992, 1993) a chaining account of this behaviour is insufficient to account for the emergent performances - especially as the A and B stimuli had never occurred together in training, yet the subjects were able to respond accurately on mixed A/B probes. Instead Stromer et al believed that these emergent performances were based on the relations established between the stimuli in training and the development of functional sequence classes.

"These outcomes suggest that the subjects' baseline behaviour was based on relations among the stimuli and that these relations reflected stimulus control by the relative as well as the absolute positions of the stimuli in the baseline sequences. . . . Furthermore, the mutual substitutability of the A and B stimuli on mixed . . . probe trials support a conclusion that the subjects had learned five two-member functional classes, each consisting of an A and B stimulus that occupied the same serial position in the sequences trained directly " (p.257).

This experiment replicated and extended previous studies suggesting that ordinal learning might be established at an earlier mental age than had previously been thought. Lutkus and Trabasso's (1974) findings suggesting the establishment of ordinal relations in subjects with mental ages of 5-6 years would have predicted these results with DC and AP. However no previous research would have predicted the successful performance of BF on these tasks (PPVT age equivalent 2 years 10 months)

Summary

The different studies on transitive inference, stimulus equivalence and sequence classes show that human subjects can demonstrate a large number of emergent relations in a variety of contexts. The study by Lipkens, Hayes and Hayes (1993) showed that relations of equivalence could be demonstrated with a child of just 17 months old, and Stromer, Mackay, Cohen and Stoddard (1993) were able to demonstrate ordinal learning with a subject with learning disability who had a suggested mental age of 2 years 10 months.

The fact that human subjects are able to derive relations such as this at such an early stage suggests that these "emergent abilities" may be useful from quite early on with regard to integrating new information with existing knowledge. In fact

these skills may be more than “useful” in structuring existing knowledge and the learning of new relations, they may prove to be essential to these processes. The power of these processes may be illustrated by a study by Sidman, Kirk and Willson-Morris (1985). In this study subjects were taught 15 conditional discriminations (five groups of three) which established three six-member stimulus classes. As a result of the training given the subjects demonstrated 60 new conditional discriminations that they had not been explicitly taught. The number of relations derived in relation to the number of relations explicitly taught gives some idea of precisely how useful this ability to derive new relations from existing knowledge can be and could help explain the rapid rate of learning shown by normally able young children. While the study by Sidman, Kirk and Willson-Morris (1985) focused on derived relations in the context of stimulus equivalence, the paradigm of transitive inference similarly shows a substantial number of derived relations resulting from a small number of explicitly trained relations. For example, the procedure developed by Bryant and Trabasso (1971) explicitly trains four relations, the overlapping stimulus pairs $A > B$, $B > C$, $C > D$, $D > E$. This training can then result in the derivation of up to six new non-adjacent relations. As the number of overlapping pairs explicitly trained increases then so too does the possible number of relations which can be derived. For instance, training one further overlapping pair $E > F$ has the potential to produce another four non-adjacent relations. Given the potential power of the ability to derive these emergent relations it would be surprising if this ability were *not* important in normal learning.

The paradigm of transitive inference has been investigated mainly in the context of cognitive psychology while stimulus equivalence has been very much rooted in behaviour analytic techniques. However, the experiments on sequence classes have drawn heavily on previous research in both the areas of inference and equivalence. The fact that the results obtained can be shown to be consistent with previous results in both these areas suggests that these paradigms are not investigating separate types of learning but rather, different manifestations of processes which facilitate learning in general.

The techniques and stimuli used to investigate these paradigms of transitive inference and stimulus equivalence tend to be very arbitrary and abstract. This is necessary in order to investigate the precise conditions under which these emergent relations can be demonstrated although, it can have the result of making this research seem very academic and not particularly relevant to natural learning. It has been suggested that stimulus equivalence is relevant to symbolic skills such as

language. For example, it may establish the relations between spoken and printed words and pictures and allow the development of abilities such as reading, writing, telling the time etc. Lazar (1977) suggested that the study of sequence relations may give some insight into how relations of grammar and word order can be learnt by young children without being directly taught. So while the paradigms of transitive inference and stimulus equivalence may appear abstract and academic they may in fact give some insight into how quite complex symbolic behaviours seem to be established quite naturally with normally developing human subjects.

Table 4.1. Training Relations for groups G0, G1 and G3, Siemann and Delius (1994)

Plus and minus signs indicate reward and punishment. Boldface type indicates pairs inferring a linear A to E hierarchy in transitive inference tasks. Italic type indicates pairs inconsistent with this linear hierarchy. (Adapted from Siemann and Delius 1994)

	A	B	C	D	E
Group G0					
A		A+B-	A+C-	A+D-	A+E-
B			B+C-	B+D-	B+E-
C				C+D-	C+E-
D					D+E-
E					
Group G1					
A		A+B-	A+C-	A+D-	
B			B+C-	B+D-	B+E-
C				C+D-	C+E-
D					D+E-
E	<i>E+A-</i>				
Group G3					
A		A+B-	A+C-		
B			B+C-	B+D-	
C				C+D-	C+E-
D	<i>D+A-</i>				D+E-
E	<i>E+A-</i>	<i>E+B-</i>			

Fig. 4:1. Graphic representation of the task structures G0, G1 and G3. The arrowheads represent inequalities not directions.
From Siemann and Delius (1994)

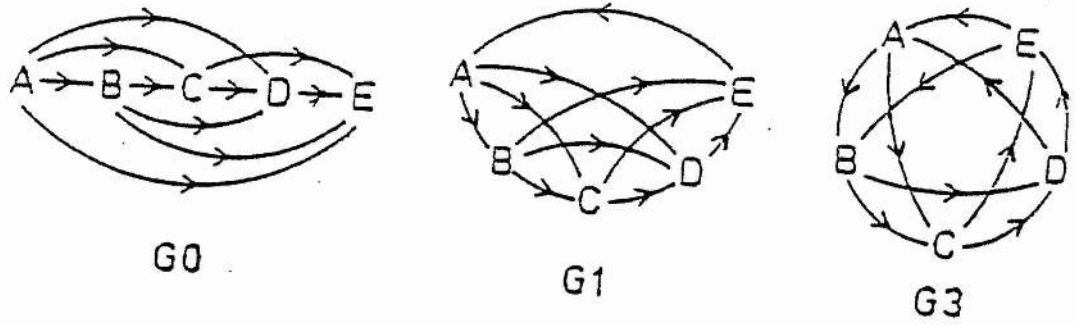
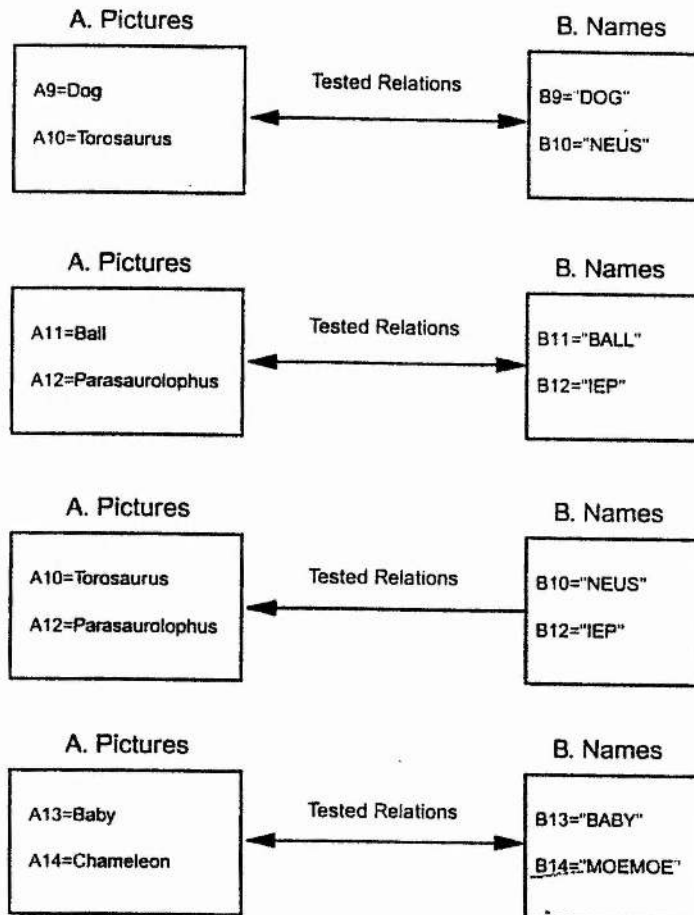


Fig. 4.2. The tested known (known name-picture and known picture-name) relations and derived (novel name-novel picture and novel picture-novel name) relations in Experiment 4. (A, pictures; B, names).

From Lipkens, Hayes and Hayes (1993).



CHAPTER 5

The Basis of Stimulus Equivalence

The defining feature of stimulus equivalence is the emergence of new relations from previously established relations, in the absence of direct training or reinforcement. The biggest problem in equivalence research has been how to account for the emergence of these untaught relations. Just what processes underlie the derivation of these untaught relations is not yet clear. There have been three main explanations proposed to account for their emergence. Dugdale and Lowe have suggested that equivalence is the result of verbal mediation; that language leads to equivalence. Sidman suggests that equivalence may be a primitive stimulus function; that like reinforcement, discrimination etc. it may be unanalysable. Hayes has suggested a relational frame theory to account for equivalence relations; that humans' histories include training that leads to the development of generalised arbitrarily applicable relational responding.

Naming as the Basis for Stimulus Equivalence

Dugdale and Lowe (1990)

Dugdale and Lowe (1990) proposed that naming provided an explanation for the emergence of equivalence relations. They suggested that intuitively naming made sense as an explanation - after all five 10p coins and one 50p coin could be given the common label "fifty" and accepted as equivalent even though they are very different stimuli physically. Or possibly naming could mediate equivalence, even if the stimuli weren't given a common name, as long as the names could be incorporated in a verbal rule that linked the sample and comparison.

As evidence for the role of naming Dugdale and Lowe noted that if equivalence does require naming, then non-humans should be unable to pass equivalence tests, and certainly at the time of this paper there had been no convincing demonstration of equivalence with non-humans. This difficulty in demonstrating equivalence relations with non-humans did seem to implicate language as being important in equivalence relations and thus naming would be a plausible basis for equivalence class formation. Dugdale and Lowe then asked, if language was the crucial factor in the development of equivalence relations, would non-human subjects who had been given language training be able to demonstrate equivalence relations. To test this, one of the authors (Dugdale) ran a series of tests with three chimpanzee subjects who had all participated in an ape-language training program conducted

by Duane Rumbaugh and Sue Savage-Rumbaugh. These chimps were all able to communicate by touching lexigrams associated with objects, actions or locations which were arranged on a keyboard. However during this training and testing for equivalence the chimps did not have access to a lexigram keyboard.

The aim of the experiment was to teach the chimps two A-B matching relations and then give B-A probes to test for the emergence of symmetry. The A stimuli were two forms (a Y shape and a zigzag) and the B stimuli were two colours (green and red). Given the Y sample the chimps were taught to select the green comparison and given the zigzag sample to select the red comparison (see fig.1). Two of the three chimps (Sherman and Lana) learned the AB matching task, the third did not. When the subjects reached criterion on this AB task, the rate of reinforcement was lowered so that only 1 in 5 correct responses were reinforced, and unreinforced BA test trials were then interspersed among the sparsely reinforced AB trials.

When tested, Lana's BA performance was at chance level even though her AB performance was around 90% correct. Lana had therefore failed the symmetry test, there was no evidence that the sample and comparison stimuli had become equivalent. The training and test procedure was then modified in two ways to try and produce equivalence responding. Reinforcement was made available for correct responses on both baseline and test trials as it was possible Lana had perceived that reinforcement was never given on test trials and thus did not attend to the stimuli on BA trials. Secondly, identity matching trials were added to the baseline so that Lana had experience with all stimuli functioning as both samples and comparisons. Despite these modifications Lana's performance on test trials remained at chance level. The other chimpanzee, Sherman, who had completed AB training was also tested on the BA symmetry relation. In all his test trials Sherman received reinforced test trials and identity matching controls as had been introduced with Lana. Despite this Sherman's performance remained at around chance level on symmetry tests.

Both chimps then failed the symmetry tests given after training a single arbitrary matching relation. As symmetry is a necessary property of equivalence, failure on this task was sufficient to conclude that the stimuli had not become equivalent. Dugdale and Lowe suggested that these chimpanzee subjects must have been among the favourites to be the first non-humans to pass a symmetry test. As chimpanzees they were one of the closest relations to humans in the animal

kingdom, and even more importantly these particular subjects had an unprecedented history of complex training and testing. Yet despite this, and several modifications to the test procedure designed to maximise the chimpanzees chances of success on the symmetry test, they showed no evidence of having derived the BA relations as a result of the training given on the AB conditional discrimination.

Negative evidence of this sort provided some support for Dugdale and Lowe's theory that non-humans would be unable to pass equivalence tests because stimulus equivalence was dependent on linguistic ability. Dugdale and Lowe felt that more direct evidence of the importance of naming in equivalence came from studies such as Devany, Hayes and Nelson (1986, see chapter 4). where normally able 2-year olds and 2-4 year old learning disabled children with functional spontaneous speech and signing demonstrated equivalence class formation, while 2-4 year old learning disabled children with no functional verbal skills failed to demonstrate stimulus equivalence. The study by Barnes, McCullagh and Keenan (1990, see chapter 4) would provide additional support for this view by effectively replicating Devany et al's findings but having eliminated degree of learning disability as a confounding factor. Devany et al had concluded that language and stimulus equivalence were closely related, albeit it was not clear in what way they were related.

Dugdale and Lowe felt that several studies demonstrated not only that language and equivalence were correlated, but in fact that language was necessary for equivalence. Lowe and Beasty (1987) showed that children aged younger than 4 who had initially failed equivalence tests subsequently passed these tests when taught to name the sample-comparison pairings during baseline training. The children were taught to match a vertical line sample (A) to a green comparison (B) - AB, and a vertical line sample (A) to a triangle comparison (C) - AC. The children were then tested for the emergence of BC/CB (green-triangle and vice-versa) equivalence relations. Subjects who failed the equivalence test were taught to say "up-green" on the A-B training trials and "up-triangle" on the A-C training trials. Following this training the subjects responded correctly on the BC and CB tests for equivalence. Dugdale and Lowe believed that this provided evidence that naming was necessary for the derivation of equivalence relations. The question then was *how* did naming produce these equivalence relations. One possibility suggested (Lowe, 1986) was that the triangle and the green stimuli both controlled the common word "up" spoken by the subject. This common name may have caused the triangle and the green stimuli to become equivalent.

To investigate this suggestion further, Dugdale (1988) carried out another experiment, trying to establish AB/BA symmetrical relations by means of a common name. The relation to be established and the stimuli used were the same as those trained with the chimpanzee subjects Sherman and Lana (see fig. 1). However the training used to try and establish this was very different. The subjects in this experiment were six 4 and 5 year old children who had consistently failed to learn the AB conditional discrimination relations. Five children were taught to give common labels to the stimuli. The stimuli were presented to the children one at a time, in response to the Y shape and the green colour the children were taught to say "OMNI", and in response to the zigzag and the red colour to say "DELTA". Effectively then the children were taught the labelling relations AX and BX (see fig. 2). These relations were learned rapidly and the children were then given AB and BA matching-to-sample test trials

One subject was given a different training procedure. This subject (Jessica) was taught to select the A and B stimuli conditionally upon labels dictated by the experimenter so that when the experimenter said "OMNI" Jessica selected the Y shape or the green comparison and when the experimenter said "DELTA" Jessica selected the zigzag shape or the red comparison. This effectively taught the relations XA and XB (see fig. 3). Jessica then also received AB and BA matching-to-sample tests.

Four of the five subjects taught AX and BX common labelling demonstrated appropriate AB and BA matching relations. These subjects gave the appropriate common labels to the sample from the beginning of testing. One subject who had been taught AX and BX and Jessica, the subject who had been taught XA and XB upon hearing common labels, responded at around chance level on the AB and BA tasks. Neither of these subjects applied common labels to the samples. A common labelling intervention was then introduced so that both subjects were asked "Is it OMNI or DELTA ?" when the sample appeared. Both subjects then produced appropriate common labels for the same stimuli and there was also an improvement in their AB and BA test performances so that they also responded appropriately on these matching relations. Dugdale and Lowe believed that "in the case of all six subjects, each shape and the corresponding colour had become equivalent through the subjects' use of common set X labels" (p. 129).

However Dugdale and Lowe also noted that performance on the BA task tended to be better than performance on the AB task, and that the BA relation tended to emerge prior to AB. Analysis of why this may have occurred gave some indication of how Dugdale and Lowe believe naming may produce equivalence. Dugdale and Lowe described this process in terms of the relations established between the green sample (B) and the Y shape (A), using verbal labels.

When the green sample appeared the subject must say "OMNI", relation BX which was directly established in training. Having said "OMNI" the subject then had to select the Y comparison (relation XA). Thus a two-stage process is established - BX and XA therefore BA. While BX had been directly established by training, relation XA had not (saying "OMNI" and selecting the Y shape). Dugdale and Lowe thus suggest that when relation AX was established so too was the potential for the symmetrical relation XA - the training had resulted in the formation of symmetrical relations between the A shapes and the X labels. This allowed the emergence of XA and thus accurate responding on BA.

The AB relation can be seen in a similar way. The subjects had to see the Y shape and say "OMNI" (AX) and having said "OMNI" then select green (XB). The AX relation had been directly established by training but the XB relation had not. However, while the AB relation did eventually emerge it did not do so immediately. Initially when the Y shape appeared the subject did label it "OMNI", but having said "OMNI" did not consistently select green. While the BX relation had been established by training it's XB symmetry counterpart (necessary for the AB relationship) had not been established. This may have been due to spontaneous labelling of the coloured stimuli.. All the subjects who derived BA before AB spontaneously labelled the colours with their appropriate names "RED" and "GREEN". These labels may have interfered with the emergence of the XB relation where the subjects had to say "OMNI" and then select Green. This interference may not have occurred with the emergence of the XA relation if the subject did not have names for the A stimuli prior to being the X labels.

Therefore, the BA relation required the relations BX, which had been directly taught, and the relation XA which may have emerged immediately from the taught AX relation as there were no prior names to interfere with this symmetrical relation. However the AB relation required the relation AX, which had been directly taught, and the relation XB, derivation of which may have been affected by the conventional labels "green" and "red" which the subjects displayed prior to

being taught the set X labels. This may explain why relation BA emerged prior to AB.

What is most notable about this account is that it emphasises the bidirectionality of the relations between the stimuli and the set X labels. "Given the presence of bidirectionality between words and shapes, should we say that the behaviour in this instance is best described, not merely as labelling, but as *naming*?" (P.132, italics in original)

The relations established between the stimuli and the labels proved to be symmetrical, and as symmetry appears to be one of the defining properties of symbolic behaviour, (Catania 1986, Devany et al 1986) Dugdale and Lowe proposed that "naming is itself a symbolic skill that involves bidirectionality" (p.132) and this skill of naming would require "not only should a particular stimulus control a subjects' verbal response, but the subjects' verbal response should also exert control over other behaviour (e.g. selection) with respect to that particular stimulus" (p.133). Naming then requires both language production and comprehension, and most importantly, these skills are not independent but linked within a single "emergent" symmetrical relation. This is why the children's behaviour in this experiment can be described as symbolic and "naming"; having learned to produce a verbal response conditional upon a stimulus they were able to produce the reverse without further training. If both components had been trained then they may not have been displaying true naming.

Dugdale and Lowe felt that the data from the subjects in the experiment above supported this interpretation. Initially the labels "OMNI" and "DELTA" were not symmetrically related to the set B colours, hence the sub-criterion performance on the AB tests. However, according to Dugdale and Lowe's interpretation the subjects were able to name BOTH the colours and the shapes with the labels "OMNI" and "DELTA" and thus the stimuli became fully equivalent.

The question then was how do the data from the chimpanzee subjects fit such an interpretation of equivalence, especially as these subjects had received some linguistic training. Some children also failed tests for equivalence initially but subsequently passed equivalence tests once they could apply a common label to each member of a stimulus set. According to Dugdale and Lowe these children's behaviour was symbolic at that point and that in using the common labels they were in fact naming the stimuli. It was possible then that the chimps were not

“naming” the stimuli in the sense described here and their lexigram responses may not be functionally equivalent to the naming shown by the children.

One problem with this account of naming as the basis of equivalence, which Dugdale and Lowe acknowledge, is that they are arguing not only that naming is necessary for forming a symmetrical relation, but also that naming itself is a symmetrical relation. Dugdale and Lowe get round this problem by differentiating between stimulus-response symmetry (naming) and stimulus-stimulus symmetry (an equivalence relation). The stimulus-response (naming) symmetry is primary and they argue that this underlies stimulus-stimulus symmetry and equivalence relations.

Dugdale and Lowe suggested that naming may arise as a result of the natural training within a developing child’s environment. Within this environment a child learns to function as both a speaker and a listener, and also learns to say a particular word conditional upon a stimulus and also to produce that stimulus conditional upon the spoken word. Thus, the naming relation may be rooted in the verbal community in which many exemplars of stimulus-response reversals are reinforced.

Dugdale and Lowe had suggested that there was evidence to show “not only that equivalence is absent when naming is absent but also that equivalence emerges when naming is introduced” (p. 124). Naming itself, they argued, was a bidirectional symbolic skill, a stimulus-response relation which then resulted in the emergence of bidirectional stimulus-stimulus relations. They believed that the emergence of equivalence relations required appropriate naming behaviour.

Eikeseth and Smith (1992)

A study by Eikeseth and Smith (1992) also studied the possible role of naming in the formation of equivalence relations. This study carried out a number of manipulations to try and investigate if, and how, naming facilitated equivalence. The training and testing in this experiment took place in 5 phases, arranged in an ABCDA design:

Phase 1 - A - Two sets of conditional relations were taught and equivalence tests given, without any programmed naming.

Phase 2 - B - If the subjects did not display equivalence in phase 1, naming was added to the conditional discriminations and equivalence was retested.

Phase 3 - C - This assessed the effects of naming the stimuli without giving match-to-sample training.

Phase 4 - D - New, unnamed stimuli were related to the named stimuli trained in phase 3 by matching-to-sample. Would untrained relations develop between the named stimuli and the new unnamed stimuli, and would untrained relations develop between the new unnamed stimuli ?

Phase 5 - A - Again, conditional relations were taught and equivalence tests given, in the absence of programmed naming. Would the subjects still fail to develop equivalence relations ?

The subjects in this experiment were four high-functioning, pre-school aged autistic children. The authors felt that these subjects were an appropriate group to study because some research suggested that autistic children seemed less likely to display untrained relations than other human populations (Lovaas 1977). This meant that these subjects might give a good indication of the conditions under which equivalence relations would or would not develop. Prior to training the subjects were tested on a number of measures e.g. the Leiter International performance Scale and the Peabody Picture Vocabulary Test - Revised. The subjects ranged in chronological age from 3.5 to 5.5 years. In general all subjects displayed average or above average visual-spatial skills but displayed deficits in language.

In phase 1 of the study, the subjects were taught AB and AC baseline conditional discriminations using a match-to-sample procedure. Having mastered these relations the subjects were tested on the BC and CB equivalence relations. Criterion was set at 90% correct responding. If the subjects failed to demonstrate equivalence, symmetry was tested by probing performance on BA and CA relations. Regardless of the result on this test the BC and CB relations were probed again. Finally the AB and AC relations were tested to check for any extinction of the baseline discriminations. All four subjects learned the AB and AC conditional discriminations and, having reached criterion on this task, maintained these relations throughout phase 1. However none of the subjects demonstrated the emergent BC and CB conditional relations. On the BA and CA symmetry relations, two subjects displayed accurate performance, one subject performed at chance level, and one subject performed at above chance level but below criterion.

Phase 2 tested whether the introduction of naming would facilitate the emergence of the BC and CB equivalence relations. First the subjects were taught to speak one

common name when shown each member of the first stimulus class (A1, B1, C1), and to speak another common name when shown each member of the second stimulus class (A2, B2, C2). The subjects then had to perform the AB and AC baseline discriminations while correctly labelling the stimuli. Having mastered this task the BC and CB relations were reassessed. Two subjects (Joe and Danny) immediately demonstrated the emergence of the untrained BC and CB relations. A third subject (Trey) performed at above chance level but below mastery. The fourth subject (Rory) continued to perform at chance level. For two of the subjects then, naming appeared to facilitate the emergence of equivalence relations.

In phase 3 the subjects were presented with new stimuli D, E and F. They were taught to speak one common name in response to the first potential stimulus class, D1, E1 and F1, and another common name in response to the second potential stimulus class, D2, E2 and F2. Once the subjects could label the stimuli correctly they were tested on the DE and DF conditional relations to see if naming itself was sufficient to establish these relations in the absence of matching-to-sample training. The EF and FE relations were then assessed, but any subjects who had failed to develop the DE and DF relations were taught these relations before EF and FE were probed. After learning to label the stimuli, two subjects (Trey and Danny) immediately showed mastery of the DE, DF, EF and FE relations. The other two subjects did not demonstrate these relations and were explicitly taught the DE and DF conditional discriminations. Following this training Joe demonstrated the emergent EF and FE relations while Rory performed at above chance level but below criterion. In phase 3 a visual-naming procedure had been introduced for Rory as this subject had lower levels of language skills than the other subjects but had good visual-spatial skills. The visual names were block structures built on top of the stimuli e.g. the "name" for the first stimulus class was a tower of two blue blocks, and the "name" for the second stimulus class was a pyramid of three orange blocks. Using this procedure Rory's performance was not perfect, but performance was still better than with the vocal naming procedure. Overall then, Eikeseth and Smith felt that the naming procedure seemed to facilitate the emergence of untrained conditional relations, particularly for Trey and Danny but also to some extent with the other two subjects.

In phase 4 two new, unnamed stimuli were related to each of the classes of D, E and F stimuli by teaching EG and FH conditional discriminations (i.e. E1G1, E2G2, F1H1, F2H2). Would untrained conditional relations develop between the named stimuli and the new, unnamed stimuli? To test this, performance on the GD

and HD conditional discriminations was assessed. Also, would untrained conditional relations develop among the new unnamed stimuli? To test this performance was assessed on the GH and HG conditional discriminations. One subject (Trey) demonstrated all four emergent relations, GD, HD, GH and HG, indicating that relations had formed between the named stimuli and the new, unnamed stimuli, and also between the new, unnamed stimuli. The results were rather mixed for the other subjects. Joe performed accurately on GD and HD only, indicating the relations had formed between the named stimuli and the new, unnamed stimuli, but not between the new unnamed stimuli. Danny demonstrated emergent relations between the new unnamed stimuli on GH and HG and to some extent between the named stimuli and the new unnamed stimuli by accurate responding on HD but not on GD. Rory showed accurate performance only on HD (new unnamed - named) but did approach criterion on GH and HG.

In phase 5, the subjects were once again taught two conditional relations and tested for the emergence of equivalence relations, in the absence of programmed naming. With a new set of stimuli the subjects were taught the conditional discriminations IJ and IK using a matching-to-sample procedure and then tested for the emergence of JK, KJ, JI and KI. Two subjects, Joe and Danny demonstrated the emergence of all these relations. The other two subjects, Trey and Rory, did not demonstrate any of these relations.

During phases 2 and 3 the subjects were taught to label the stimuli within each potential class with a common name. During the match-to-sample probes in phase 2, 3 and 4, three subjects (Trey, Joe and Rory) continued to label the stimuli correctly even though naming was not requested on these probes. The subject labelled the individual stimuli correctly even when they responded incorrectly on match-to-sample tests. One subject (Danny) did not use overt labels on the match-to-sample tasks and labelled the stimuli only when required. None of the subjects used overt labels during phases 1 and 5 when naming was not directly programmed for any of the stimuli.

Eikeseth and Smith felt that these results did show a causal relationship between naming and the development of equivalence relations. One subject, Trey, consistently responded accurately on probes of untrained relations when naming was programmed, and equally consistently failed to show emergent relations when naming was not programmed. While the other subjects performances were not as

consistent as Trey's in this respect, they were also more likely to display untrained conditional relations when naming was programmed.

Eikeseth and Smith suggested that naming established the different stimuli that had been given the same name as members of a functional class (Goldiamond, 1962), i.e. the different stimuli came to evoke the same response - saying a certain name. Thus, even though the functional classes may have been established outside a match-to-sample task, emergent relations were demonstrated within a match-to-sample context. Eikeseth and Smith thus believed that "the establishment of functional classes through naming facilitates the development of untrained conditional relations, some of which appear to be equivalence relations" (p.132). However it is unlikely that the relationship between naming, functional classes and equivalence classes is a simple one. After all, one subject (Danny) clearly demonstrated equivalence relations in phase 5 when no naming was programmed, and Danny only used overt names for the stimuli when required to do so. So while it is not possible to exclude the possibility that Danny was naming the stimuli covertly, it is equally possible that he was not naming the stimuli at all, in which case naming would not provide a sufficient explanation for the emergence of the untaught relations. Eikeseth and Smith themselves acknowledge that the relationship between functional and equivalence classes is likely to be complex. A study by Sidman, Wynne, Maguire and Barnes (1989) concluded that while there seems to be a relation between functional and equivalence classes, they do not constitute the same behavioural process. However, Murray Sidman has since substantially re-evaluated this conclusion (cf. Sidman, 1994; discussed more fully later in this chapter)

However this study by Eikeseth and Smith was a useful attempt to investigate a possible causal link between naming and the emergence of stimulus equivalence, and to determine whether naming in itself is sufficient to establish equivalence relations. The conclusions from this study would seem to be that naming may facilitate equivalence class formation but it is not clear that naming causes equivalence class formation.

Horne and Lowe (in press)

A recent study however, Horne and Lowe (in press) has provided a very detailed account of precisely how the authors believe naming brings about emergent behaviour, such as stimulus equivalence. According to Horne and Lowe "naming" is the basic unit of verbal behaviour and it is this which, when learnt, plays a

crucial role in the development of stimulus classes and of symbolic behaviour. This account of naming and symbolic behaviour draws heavily on B.F. Skinner's analysis of verbal behaviour (1957) and the basic verbal relations which he identified (i.e. the "tact", the "mand" and the "intraverbal") are all seen as becoming variants of the basic name relation. Horne and Lowe believe that the development of naming requires that conventional listener and speaker behaviours combine and that this occurs as a result of echoic responding, either overt or covert.

This study focuses on the way verbal behaviour may develop in a very young child as a result of behavioural interactions between the child and its environment/primary caregivers. In particular they describe how they believe naming develops from earlier, pre-linguistic behaviour. It has been suggested that before children learn to speak they must learn to listen and Skinner (1957) termed this "listener behaviour". Horne and Lowe felt that the development of this "listener behaviour" was crucial to the development of linguistic behaviour and described the circumstances under which it might be established. Ultimately this behaviour should become discriminative for object or event-related behaviour. First the caregiver would produce a vocal stimulus such as an object name in the presence of both the object and the child. Typically the caregiver would observe what the child was looking at, and then indicate that object or event before naming it. This "joint regard" increases the likelihood that the name produced by the caregiver will become discriminative for the child looking at the object which the caregiver has indicated. This may then develop when the child points to an object or gives an object to the caregiver, and this behaviour becomes discriminative for the caregiver naming that object. At the same time the caregiver may teach the child how to perform conventional behaviour in relation to the object. The child's imitation of these conventional behaviours is reinforced and this imitation may be prompted or shaped by the caregiver. As this behaviour is established, the caregiver's vocal stimulus will increasingly precede the child's behaviour rather than accompanying it. In this way the caregiver's vocal stimulus becomes discriminative for the child performing these object-related behaviours until the child reliably behaves appropriately to that object given the vocal stimulus. Thus the child can be said to have acquired listener behaviour.

Horne and Lowe illustrated this with the example of a caregiver seeing a shoe, pointing at it, and saying to the child "where's shoe?". By the process described above the child may learn to orient to the shoe, and point to it or hand it to the

caregiver. This listener behaviour is extended when the child learns that the term given to that object may also apply to a number of other objects. For example, the term "shoe" is not applied just to one particular shoe but rather to a variety of different shoes, and in this way a class of objects may be established by the caregiver naming each of the different exemplars "shoe". Horne and Lowe state that in this way the child learns a series of relations between the auditory stimuli produced by others and her own responding to a class of stimuli. This listener behaviour may be extended following several examples of vocal stimuli such as "where's the shoe ?", "where's the cup ?" etc. The first "Where's the ...?" part of the stimulus may come to elicit orienting behaviour, such as pointing, while the second part of the stimulus determines the object to which this orienting occurs. New listener behaviours can then be added to the child's repertoire by teaching behaviours such as "Pick up . . . shoe/spoon etc.". The "pick up . . ." behaviour may thus generalise to all the objects to which the child has learned to orient.

The crucial feature of the acquired listener behaviour is that it is determined by other's speaking. For Horne and Lowe's theory it is important that the child progresses to being a speaker/listener and they suggest that this occurs through echoic responding. One of the precursors of echoic behaviour may be the consonant-vowel "babbling" vocalisations displayed by human infants. There are several findings to suggest that these vocalisations are dependent on exposure to the speech sounds of a verbal community. Horne and Lowe suggest that raising the operant level of these vocalisations is an important step towards the acquisition of speech. The "echoic" is one of the classes of functional relations in verbal behaviour defined by Skinner, and involves responding or "echoing" a heard verbal stimulus. Horne and Lowe believe that this echoic relation is important in establishing the second link of the naming relation - a relation between a child making an utterance and hearing that utterance. The caregiver sees an object and names it, creating an auditory stimulus which the child attempts to reproduce, and is reinforced for doing this by the caregiver. However the caregiver's initial vocal stimulus also occasions the child orienting to the object (as seen previously). The child's echoic reproduction of this vocal stimulus creates another auditory stimulus, functionally equivalent to that of the caregiver, and so itself causes listener behaviour (orienting) to the named object. This may produce further echoic repetitions of the vocal stimulus. Horne and Lowe believe that it is at this point that the child becomes a speaker-listener with respect to her own verbal stimulus. It has been suggested that the child's repeated echoing of the object name may sustain the listener behaviour of orienting to, and interacting with, the object. Initially the

echoic responses must occur overtly in order for them to be reinforced by the caregiver. But it is possible that this echoic behaviour will become increasingly covert and "abbreviated", particularly in the case of repetition or self-echoing. The listener behaviour occasioned by these covert echoics will reinforce this echoing and serve to maintain it, so that once the child has learned generalised echoic behaviour, and to emit formerly overt echoics covertly, new echoics may be acquired solely at a covert level. Although the child's speaker-listener behaviour is still only instigated by other's vocalisations, Horne and Lowe suggest that this echoic behaviour sets the conditions for the object itself to directly control the child's verbal behaviour and in this way produce naming.

Important to this development of naming is the verbal relation the "tact". This is a verbal relation that results from, or is linked with, that which has gone before. The tact may come about as a result of repeated echoic interactions. For instance, the caregiver may give the child an object and if the child says the name of that object, the caregiver may reinforce this responding by smiling, praising the child etc. So that eventually that object becomes discriminative for the child producing it's name. Initially the caregiver may present the object and say the name, e.g. "shoe". This verbal stimulus causes the child to orient to the object and echo it's name until eventually the child will produce the appropriate verbal response without any need for the caregiver to speak. Horne and Lowe suggest that it is at this point, when the object alone occasions the verbal response and accompanying listener behaviour, that it can be said that the child has learned to name.

However this relation alone would also fulfil the criteria for a tact as defined by Skinner (1957) and Horne and Lowe state that the properties of naming go beyond those of tacting. "The tact relation between a stimulus and a response is uni-dimensional and non-symbolic. On the other hand . . . the name relation as outlined here has all of these defining characteristics of symbolic behaviour". These defining characteristics being that a name can *represent* a stimulus; *refer* to it; *stand* for it; and *specify* it; which a tact cannot. Horne and Lowe suggest that when a child sees an object and makes a "tact" response, this generates a verbal stimulus which causes her to re-orient to that object (as described previously). However this verbal stimulus may cause her to "see" not just that particular object but any of the other instances in her listener behaviour class. Thus, Horne and Lowe suggest, naming establishes bi-directional relations between a class of objects or events and the speaker-listener behaviour they occasion. They believe that this naming behaviour is a fusion of conventional speaker and listener functions which was

brought about by echoic behaviours. There also seems to be evidence to support this suggestion that listener behaviour is closely linked to production. For instance, Harris, Yeeles, Chasin and Oakley (1995) found evidence that once the child learns to act as a speaker-listener with respect to one object, they will then be able to name other objects in the same stimulus class to which they have previously only responded as a listener.

The name relation will also be extended beyond the original object to include new items. Horne and Lowe suggest that this extension of the name is controlled by the child's verbal community, with idiosyncratic extensions being corrected by the child's caregiver. For instance, an over-extension may be corrected - e.g. "No, that's not a dog, that's a cat" while under extensions may result in the caregiver pointing out further exemplars and establishing the name relation to these.

The authors suggest that initially joint attention to objects is established by cues such as pointing at objects and phrases such as "Where's the . . .?" etc. and several repetitions of these behaviours may be necessary to establish the name relation. Eventually though, the caregiver's naming of, and pointing at, an object may be sufficient to evoke the sequence of behaviour that makes up the name relation and the names may be learnt after the caregiver has named the object only once or twice. Thus Horne and Lowe argue that the name relation is established as a "higher-order behavioural relation", short circuiting earlier means of constructing it, and this higher order naming may become increasingly covert.

A further development of the theory is that the names provided by the verbal community also assign common class names to objects which, while different physically, have a shared cultural function - i.e. are functionally equivalent. Thus they suggest that naming establishes functional stimulus classes for the child. These culturally defined functions then become an integral part of the child's listener behaviour towards those objects so that when a new object also has these functions the child may apply the same name to that new object. This extension of the name is not directly trained or reinforced but rather occurs as a result of common listener behaviour towards that object. Horne and Lowe also suggest that naming can establish a number of levels of functional equivalence. For example, having learned that a particular object is called a "chair", the child can learn that a chair may also be called the more general name "furniture" establishing a more general level of functional equivalence.

The fact that naming may establish classes of objects is crucial to Horne and Lowe's account of how naming relates to emergent behaviour. If a new listener behaviour is established with some exemplars from a functional class, the name relationship established for that class may cause this new listener behaviour to generalise to other members of the class without the caregiver having to reinforce the target behaviour with each individual object. The child will treat each item in a similar way, as if it were functionally equivalent. Horne and Lowe believe that this is what happened in the study by Eikeseth and Smith (1992) when the autistic subjects were taught the same name for all the stimuli in a class and then subsequently the subjects were able to select the comparison stimulus that was in the same class as the sample stimulus. They felt that this apparently spontaneous behaviour was in fact listener behaviour engendered by the name.

On the basis of the account summarised here, Horne and Lowe define naming as: "A higher order bi-directional relation that (i) combines conventional speaker and listener behaviours within the individual; (ii) does not require reinforcement of both speaker and listener behaviours for each new name to be established; and (iii) relates to classes of objects and events".

This definition of naming, and the use of naming as an explanation for "emergent" relations has implications for the idea of "emergent" relations, such as those found in stimulus equivalence. According to Horne and Lowe these relations are established through naming, and as naming itself is a learned behaviour these relations are not in fact derived. Rather they feel that these relations are a learned behaviour, established through the learned behaviour of naming, the operations of which can be understood through behaviour analysis.

Another aspect of Horne and Lowe's naming hypothesis is that the role of naming in producing these emergent relations is enhanced through the development of the "intraverbal", another of the verbal units identified by Skinner (1957). This relation develops when the child begins to combine previously learned single conventional verbal responses to form response pairs etc. One of the most important features of the intraverbal is that it may establish bi-directional relations between the child's verbal responses. Again, Horne and Lowe suggest that this occurs through self-echoic repetition. For instance, one intraverbal a child might easily learn is "knife fork". Self-echoic repetition would produce "knife fork knife fork knife fork etc." so that fork may precede knife as well as vice versa, in this way establishing bi-directional relations between the behavioural components.

Another class of verbal behaviour identified by Skinner (1957) is "manding". This is behaviour that is primarily under the control of its consequences; it is a class of utterances that make demands on the hearer and are reinforced by fulfilment of the demand. There is evidence that children learn to name objects before they learn to mand them. Horne and Lowe suggest that manding comes about via naming - names are established first and then extended to mand objects and events. In this way the child can now mand any object that she can name, and these mand relations incorporate naming "with all of the speaker/listener functions and bi-directionality thereby entailed".

Horne and Lowe have provided an account of how they believe a child's behaviour with a single stimulus generalises to other physically distinct stimuli, and how that child, within particular contexts can come to respond to different objects and events as interchangeable. It is features like these that are examined in the stimulus equivalence literature. However Horne and Lowe believe that these behaviours are established via naming, which they believe is itself a learned behaviour and thus these "emergent" behaviours found in stimulus equivalence may be a consequence of the different stimuli being part of the same name relation.

One way which Horne and Lowe suggest that naming may produce equivalence classes is through intraverbal naming. They illustrate this with a description of the subjects' performances in the experiment by Lowe and Beatty (1987). The children were taught to match a vertical line sample to a green comparison (A1B1) and a horizontal line comparison to a red comparison (A2B2). Intraverbal echoing of the names of the stimuli could have produced the verbal repetitions "up green up green up green" (A1B1) and "down red down red down red" (A2B2). The subjects were then taught to match the vertical line sample to a triangle comparison (A1C1) and the horizontal line sample to a cross comparison (A2C2). This might have produced the verbal repetitions "up triangle up triangle up triangle" (A1C1) and "down cross down cross down cross" (A2C2). Thus the A1 stimulus, the vertical line, is equally likely to occasion the child saying "up green" or "up triangle", and similarly the A2 stimulus, the horizontal line, is equally likely to occasion the child saying "down red" or "down cross". These may then give rise to the three name intraverbals "up green triangle"/"up triangle green" and "down red cross"/"down cross red". As a result, any of the stimulus names may come to occasion the other two that have been linked with it, and the names become interchangeable. It is at this point Horne and Lowe suggest an intraverbal "equivalence" class has been formed and subjects will pass formal tests of equivalence within the matching

paradigm. For example, in the experiment by Lowe and Beasty (1987), having been taught the matching relations A-B and A-C the subjects would be tested on B-C and C-B tests of equivalence relations. So, given a C-B test trial, Horne and Lowe suggest that if the triangle (C1) stimulus appears, the subject will name it as "triangle", which will then occasion him/her saying "up" and "green". Pressing the sample produces the green (B1) and red (B2) stimuli as comparisons. Having heard himself or herself say "green" the subject will then orient to, and select, the green comparison rather than the red. As similar relations should hold between the other C-B and B-C stimuli this behaviour would be seen as indicating the derivation of stimulus equivalence.

Another possible way that equivalence classes could be demonstrated is through common naming. Horne and Lowe suggest that this common naming may develop as the subjects learn the conditional relations. For example, when learning A1-B1 and A1-C1 the subject may say the names "up-green" and "up-triangle". As the name "up" is common to both discriminations the intraverbal name pair may contract to a common name "up" which will control responding to all the stimuli within that stimulus class. Similarly A2-B2 and A2-C2 may produce "down-red" and "down-cross" which will contract to the common name "down", controlling responding to all the stimuli within that stimulus class. So Horne and Lowe suggest that common naming may provide another explanation of how subjects may "derive" equivalence relations as a result of naming.

This paper by Horne and Lowe has put forward an account of how naming behaviour develops as a result of interactions between a very young child and that child's principal caregivers. Horne and Lowe believe that the verbal behaviour developed in these interactions produces bi-directional speaker-listener naming behaviour. This naming behaviour then results in the development of stimulus classes in which the different stimuli become mutually substitutable as a result of the names established, whether through common naming or intraverbal naming. In this way Horne and Lowe explain the development of "emergent" relations such as those found in tests of stimulus equivalence. In fact, Horne and Lowe believe that equivalence relations are not derived but rather the direct result of learned naming behaviour. Nor, they feel should naming be seen as *mediating* equivalence class formation. Instead they believe that "naming *is* stimulus classifying behaviour" (*italics in original*).

Horne and Lowe feel that certain predictions can be made on the basis of their stated naming hypothesis and that tests of several of these predictions support naming as an explanation of equivalence class formation. The first prediction is that "because they are lacking in naming skills, non-human organisms will generally fail tests of stimulus equivalence". It is true that it has been extremely difficult to demonstrate stimulus equivalence with non-human subjects (see chapter 3), but Horne and Lowe go further, to predict that in fact there is no basis to predict that non-humans should be able to demonstrate stimulus equivalence, at least according to current accounts of non-human animal learning. Horne and Lowe suggest that the three-term (discrimination learning) contingency involved in establishing the baseline relations may in fact preclude non-human success on tests for equivalence relations. As the direction of the relation between the discriminative (sample) stimulus and the response is one way, the non-human subjects may become "locked" into this uni-directional relation with no mechanism to permit reversals to occur. According to Horne and Lowe's theory however, human subjects are freed of these constraints as they produce names which provide reversible links between their behaviour and the stimuli involved in the relation.

One problem with Horne and Lowe's theory is that Schusterman and Kastak (1993) appeared to have demonstrated equivalence relations with a sea-lion subject. However Horne and Lowe argue that this study had several distinctive features which may explain this animal's apparent success. The sample and comparison stimuli were simultaneously visible to the sea lion before it was able to respond, and also no sample response was required, only an orientation to the left or right to select one of the comparison stimuli. The incorrect comparison stimuli changed on each trial and were randomly assigned, preventing S- control by particular comparison stimuli. The subject was also given symmetry and transitivity training with a subset of the stimuli, and reinforcement was available for correct responding throughout testing. Horne and Lowe suggest that the training procedure may have established each stimulus pair (e.g. AB or BC) as a compound stimulus and eventually established responding to the outer stimulus element of the compound (i.e. responding to the comparison). Correct responding on the AC and CA relations may have been produced by "associative transitivity" as several studies have shown that when AB and BC are explicitly trained AC is present when tested. As the AC elements would have functioned as a compound stimulus, the animal responding to the outer element of the compound would produce successful performance on the CA equivalence test. Of course, this is just one interpretation of

the results obtained by Schusterman and Kastak, and equally possible is Schusterman and Kastak's explanation for the performance, that the animal was in fact demonstrating emergent equivalence relations.

Horne and Lowe's second prediction was that "Humans who are lacking the prerequisite naming skills will generally fail tests of stimulus equivalence". Unfortunately the amount of research with young children with limited naming skills on equivalence tasks is limited, but the studies by Devany et al (1986) and Barnes et al (1990) do seem to support this proposition, to some extent at least. The third prediction followed on from this and suggested that "teaching subjects particular name relations for the stimuli used in match-to-sample procedures may be a powerful determinant of subsequent performance on equivalence tests". The common naming interventions used by both Dugdale and Lowe (1990) and Eikeseth and Smith (1992) both seem to support this prediction. In both cases teaching a common name for the stimuli within an experimenter-defined equivalence class either facilitated or produced appropriate performance among the stimuli with young children. A study by Saunders, Saunders, Williams and Spradlin (1993) also found that giving names to the stimuli helped facilitate equivalence class formation with adults with mild learning disabilities, while a study by Dickens, Bentall and Smith (1993) found that the usefulness of names depended on whether those names were consistent with the experimenter defined classes trained, or contradictory to those classes.

In general then, Horne and Lowe felt that there was good evidence that appropriate naming behaviour was closely related to success on standard equivalence tests and thus supported their proposition that derived equivalence relations were a result of learned naming behaviours, acquired as a result of the fusion of speaker and listener behaviour.

Sidman's Theory of Equivalence

Sidman (1986)

Dugdale and Lowe (1990) and Horne and Lowe (in press) have provided one explanation for the emergence of equivalence relations - that equivalence is a direct result of learned verbal behaviour; naming. Murray Sidman, however, has presented a very different explanation for the derivation of stimulus equivalence. Sidman (1990) suggested that it may not be possible to derive equivalence from something more basic; that stimulus equivalence may be a fundamental stimulus

function. Sidman suggested that equivalence relations emerge at the level of a four term contingency, a suggestion first described in his 1986 paper on functional analysis of emergent verbal classes. This paper considered the basic units of behaviour analysis and how the emergence of equivalence could be described in terms of these units. These basic units form an increasingly complex systematic framework of two, three and four term contingencies.

The two-term contingency

Skinner (1935, 1938) proposed a response-stimulus relation as a basic unit for the analysis of behaviour - the two-term contingency. This contingency can be described in terms of a subject pressing a button and then receiving a coin as a consequence (see Fig.5:1). The behaviour of pressing the button, Response 1 (R1), is followed by a specific consequence, receiving a coin (C1). This relation can be described as a contingency because any other behaviour (R2) will not produce a coin. The consequence of the behaviour determines the likelihood of that behaviour being displayed again. In this case delivery of the coin increases the probability that the coin pressing behaviour will be displayed again. Any other behaviour is less likely to be displayed in this situation as it will fail to produce the specified consequence. This contingency explains how behaviour can be controlled by its consequences, which after all are events that *follow* the behaviour - the likelihood of a behaviour being displayed is determined by past contingencies.

The two-term contingency "specifies a causal relation between behaviour and environment" (Sidman, 1986, p.127) and as such is very useful in identifying positive and negative reinforcers and the effect they will have on behaviour. However this contingency is insufficient to describe ordered behaviour as the two-term contingency only describes probabilities which are active at any given moment. A more detailed contingency is required to describe more complex behaviour.

The three-term contingency.

Sidman refers to the three-term contingency as "the analytic unit that emerges when one considers the two-term contingency in relation to the changing environment" (p.219). The two-term contingency is brought under the control of a third element, stimulus S1 (see Fig. 5:2). Continuing the example given, the behaviour of pressing a button will now only produce a coin if the button has an appropriate shape on it (e.g. a square). Pressing a button with any other shape (e.g.

a circle) or any other behaviour other than pressing a button will fail to produce a coin.

Skinner (1938) pointed out that this does not mean that the square now elicits the response. Rather, the two-term relation has been brought under *discriminative* control so that the square elicits the two-term contingency, increasing the likelihood that the subject will press the button to produce a coin. In this way, the environment establishes control over the two-term contingency by selectively altering the probability that the behaviour in the two-term contingency will be displayed. Thus Sidman refers to the three-term contingency as “the fundamental unit of *stimulus control*” (p. 219, italics in original).

The three-term contingency introduces the process of conditioned reinforcement to behaviour analysis. The third term S1 is the initiating component of that three term contingency. Therefore a response which produces S1 as a consequence will then activate the three-term unit initiated by S1. Activating that three term unit in itself acts as a reinforcer and so S1 will now alter the future probability of the behaviour that precedes it. Sidman suggests that the three-term contingency is also the basic analytic unit of cognition. He suggests that knowledge is inferred from observations of stimulus control and through three-term contingencies behaviour becomes related to both consequences and antecedents.

However, again, the three-term unit is not sufficient to fully describe complex behaviour. It is also necessary to be able to account for the environment’s ability to select the particular three-term units of stimulus control that are active at any moment.

The four-term contingency

The addition of one further term places the three-term contingency under stimulus control. Continuing the example given, the subject still has two buttons available, with a square on one and a circle on the other, but a third button is now added, which is sometimes green (S3) and sometimes red (S4) (see Fig. 5:3). Now the behaviour of pressing the button with a square will only produce a reinforcer if the new button is red. This establishes a four-term contingency; the three-term relation is true only in the presence of the green button (S3), not any other colour. Thus, the three term contingency is now under conditional control. The conditional control of the four-term unit is a different stimulus function to the discriminative control of the three term unit. In the three-term contingency the discriminative stimulus (S1)

is identified by reference to a differential response - the subject presses the button in the presence of the square but not in its absence. In the four term contingency the conditional stimulus does not require additional differential behaviour - even when the square is on the button, the subject presses it only in the presence of the green hue; the green hue activates the three-term contingency. There is no intervening response between S3 and S1.

The significance of the four-term unit is that the discriminative three-term unit is under contextual control, "That is how the environment establishes higher-order priorities, selectively altering the probability of three-term relations that exist in one's repertoire" (p. 225).

The four term contingency detailed in Fig. 5:3 is somewhat unbalanced in that S3 provides a context in which the subject can press the square to receive a coin but in the context of S4 no behaviour is effective. The four-term contingency shown in Fig. 5:4 has had this imbalance corrected in that now in the presence of S4 (red), pressing the button (R1) with the circle on (S2) will produce a coin (C1).

Sidman has suggested that it is at this level, the four-term conditional discrimination, that equivalence relations are established. The conditional discriminations established in baseline training in equivalence are the same in structure as those detailed in fig. 4. For instance, in standard A-B match-to-sample training S3 would be sample stimulus A1, and S4 would be sample stimulus A2. S1 and S2 would be comparison stimuli B1 and B2 respectively. In the case of equivalence, new four-term conditional relations are established without direct training although the subject has never experienced the new contingencies before.

Sidman has pointed out that when equivalence relations are established it permits us to regard the stimuli as having the same meaning as each other.

"In this way semantic correspondence emerges from non-linguistic "if . . . then" relations. Expanding the analytic unit from three to four terms establishes the potential for verbal classes to emerge" (p.233).

One of the most important aspects of the derivation of equivalence in this way is that it provides an explanation for the establishment of simple semantic correspondences without having to provide an explicit reinforcement history for each case.

However, while Sidman argues that stimulus equivalence emerges at the level of the four-term contingency, in the analysis of basic units of behaviour, the four-term unit is not sufficient to explain fully some of the more complex behaviours.

“By itself, the four-term contingency provides only a unit for describing the contextual control of three-term contingencies, a level of analysis that does not encompass the role of context in determining semantic correspondence. . . . By placing four-term contingencies themselves under environmental constraint, however, we can bring the emergence of meaning itself under contextual control” (p. 237).

By adding a fifth term it is possible to describe the environment’s selection of conditional discriminations.

The five-term contingency

In the four-term contingency described in fig. 4, in the presence of the green hue pressing the button with a square produced a coin, while in the presence of the red light pressing the button with a circle produced a coin. In the five-term contingency this conditional discrimination is itself brought under conditional control by the addition of two further stimuli S5 and S6. In the presence of S5 (tone 1) the conditional discrimination outlined above remains true. However, in the presence of S6 (tone 2), these contingencies are reversed so that in the presence of the green hue, pressing the button with a circle now produces a coin while in the presence of the red hue pressing the button with a square produces a coin. Sidman states that as the original conditional discrimination is now under the conditional control of the tones; this five-term contingency is the unit of *second-order conditional control*.

This unit of analysis may help describe how the equivalence relations described by the four-term unit function in a natural setting. In the four-term unit the equivalence relations established would be “green with square” and “red with circle”. The five term unit accounts for contextual control of these relations, in that in the context of tone 1, these stated relations hold true, but in the context of tone 2 the equivalence relations would be “green with circle” and “red with square”. Therefore the context of tones 1 and 2 shifts elements of the environment from class to class and the stimuli can be members of several different equivalence classes, the class operating at any time being determined by context.

However, one problem with this account of five-term contingencies is describing the relation between the second-order conditional stimuli (the tone) and the other stimuli. For instance, it might be possible for tone 1 to become an element of two

three-member equivalence classes - tone 1, green and square and tone 1, red and circle. As tone 1 would be common to both classes, this would mean that the green and red hues and the square and circle forms would also be equivalent to each other, thus wiping out both conditional and discriminative control of the relations. Tone 2 would also be a member of two three-member classes, causing these classes to collapse into one large class. By this both tones would now be equivalent members of one large class, thus destroying the entire five-term contingency.

However, if the equivalence classes do not increase in size from two to three members in this way, then the problem does not arise. If the second order conditional stimuli, the tones, do not enter into equivalence relations with the other stimuli but instead maintain their conditional relations, then the contextual control over the units can be maintained without destroying the relations in the units. Thus Sidman seems to be suggesting that equivalence relations are established at the level of the four-term contingency, while the five term contingency exerts contextual control over three relations, selecting those which are appropriate in any given context.

Sidman (1990)

While this paper by Sidman (1986) considered the behavioural level at which equivalence relations emerged, a later paper, Sidman (1990) considered in more detail exactly *how* equivalence comes about. The two, three, and four term units of analysis described the stimulus functions of reinforcement, discrimination, conditioned reinforcement and conditional discrimination which represent "unanalysable primitives"; they cannot be derived from something more basic. Sidman suggested that equivalence might represent another of these unanalysable primitives; that it is a fundamental stimulus function.

One justification for this statement was that, as described previously, Sidman suggested that equivalence appears at the level of the four-term contingency. The four-term unit is not just a larger three-term relation, instead the two, three and four-term relations are all fundamentally different and as the unanalysable functions or reinforcement etc. appear at different levels of analysis, the function of equivalence, appearing at the four-term level is not derivable from smaller contingencies but represents an unanalysable primitive in itself.

A second justification for Sidman's suggestion that equivalence represents a fundamental stimulus function was that he did not believe there was any evidence

to show that equivalence could be derived from anything more basic. Logic might seem a reasonable basis for the derivation of equivalence relations, but Sidman showed that there was no logical necessity for relations between stimuli to demonstrate properties of reflexivity, symmetry or transitivity. "Logic does not demand that conditional relations be equivalence relations" (p. 105).

Similarly, Sidman was not convinced that verbal behaviour could be shown to be the cause of equivalence. He agreed that common naming was likely to facilitate equivalence class development but this does not mean that naming is *necessary* for the development of equivalence relations. There may also be a problem when a subject gives the same name to each member of an equivalence class when that subject has not explicitly been taught to do so. In this case Sidman suggests that equivalence classes may have produced the common names rather than naming producing equivalence classes. Another possibility was that equivalence might be a result of verbal rules, but again Sidman suggested it was not clear whether the verbal rule or the equivalence relation came first.

Sidman suggested that possibly the most effective way to disprove the suggestion that equivalence was a result of verbal behaviour was an unequivocal demonstration of stimulus equivalence with non-human subjects. At that time (1990) there had been no such demonstration of equivalence relations, but if Schusterman and Kastak's subsequent experiment with a sea-lion subject is accepted as a demonstration of equivalence, then this would suggest that Sidman's assertion that equivalence was *not* a product of verbal rules was correct.

Sidman (1994)

In his 1994 book reviewing research on equivalence Sidman gave a comprehensive re-appraisal of his earlier suggestions that equivalence emerged at the level of the four-term unit and that equivalence could be viewed as a basic stimulus function. In this reappraisal he stated that he now felt that defining equivalence at the level of the four-term contingency was actually too restrictive and that in fact there was good evidence that three-term contingencies, simple rather than conditional discriminations, could in fact establish the pre-requisites for equivalence relations. However while his beliefs about equivalence and the four-term unit had changed somewhat, he also stated that he now believed even more strongly that equivalence was a basic stimulus function that could not be derived from anything more fundamental.

Some developments in equivalence research persuaded Sidman that an expanded notion of the relations between equivalence and contingency defined units might be more appropriate than the previous formulation of equivalence at the level of the four-term contingency. One such discovery was the finding that an equivalence relation could include not only the unit's conditional (sample) and discriminative (comparison) stimuli but also the reinforcing stimulus (p. 368). This was shown using outcome-specific reinforcement contingencies. In this paradigm the defined response produces different reinforcers. For instance, in the presence of sample A1, touching comparison B1 will produce reinforcer SR1, while in the presence of sample A2, touching comparison B2 produces reinforcer SR2. Now on test trials the stimuli that had previously served to reinforce could also function as either samples or comparisons. In tests of this sort where SR1 or SR2 function as samples and the comparisons are A1 and A2 or B1 and B2, it has been found that subjects will match SR1 to A1 and B1 and SR2 to A2 and B2, thus suggesting that equivalence class membership also extends to the reinforcers (p.369)

With this evidence that equivalence relations include three of the four elements in the four-term contingency, the next question was whether equivalence relations also include the fourth element, the response. However, as Sidman has pointed out, it is likely to be very difficult to isolate the response for testing.

Sidman suggested a possible experiment that might evaluate if the response is involved in equivalence relations. This involves training another conditional discrimination (AB), but this time there is only one reinforcer SR1 but two defined responses to the comparison stimuli R1 and R2 (see Fig. 5:5, top section). These responses are two distinct patterns of pressing. When the sample stimuli A1 or A2 are presented a single touch (R3) produces the comparison stimuli B1 and B2. In the presence of sample A1 the subject has to press comparison B1 using response R1 to produce a reinforcer, and in the presence of sample A2 the subject must touch comparison B2 using response R2 to produce the reinforcer. No reinforcement is delivered if the subject selects the wrong comparison or produces the wrong response. The subject is then taught a second conditional discrimination (BC) in the same way. This time the B stimuli function as samples, and so require only a single touch (R3) to produce the comparison stimuli C1 and C2. Correct responding to the comparison stimuli requires response R1 to C1 and response R2 to C2 (see Fig. 5:5, middle section).

Having trained relations AB and BC, CA constitutes the test for equivalence. In this case the C stimuli function as samples requiring a single touch R3 to produce the comparison stimuli. The question then is whether the subject will select stimulus A1 using response R1 and stimulus A2 using response R2. If the subject does respond to the stimuli in this way it would suggest that R1 belonged to the A1B1C1 class and R2 to the A2B2C2 class (see Fig.5:5, bottom section). However, as Sidman pointed out, a more decisive test for the inclusion of responses in equivalence classes would be one where the responses themselves could function as samples. It was the difficulty of establishing a test like this that suggested there might be a closer connection between reinforcement and equivalence than Sidman had previously considered.

One way in which it might be possible to test the relation of the response to the equivalence classes might be to establish the AB conditional discrimination with two different responses as just described. Then the subject might be taught a simple discrimination - two three-term relations. In this case stimulus D1 is present on some trials and stimulus D2 on others. When D1 is present, response R1 produces the reinforcer and when D2 is present R2 produces the reinforcer. The aim of this three-term relation is to establish a situation where R1 and R2 will occur reliably and thus can function as samples. Having trained the AB relation, if R1 and R2 have become members of the A1B1 and A2B2 classes, then on trials where R1 is sample the subject should select A1 and where R2 is sample, the subject should select A2. Thus, one test trial might involve presentation of D1, which should occasion response R1 which would produce A1 and A2 as comparison stimuli. If R1 has become related by equivalence to A1 and B1 the subject should then select stimulus A1 using response R1. A similar pattern of responding would be expected with D2, R2 and A2.

However, the problem with this experiment is being sure that selecting comparison A1 or A2 is actually under the control of responses R1 or R2 rather than stimuli D1 or D2, even though the D and A stimuli have never been presented together before. The reason for this uncertainty stems from Sidman's re-evaluation of the relationship between equivalence relations and the units of analysis. This was because several experiments had suggested that three-term contingencies could in fact establish the pre-requisites for equivalence relations (discussed in more detail later). This has implications for the experiment described above as it suggests that

“when two or more discriminative stimuli control the same two-term contingencies (the same defined response and reinforcer) those stimuli can be shown to be related by equivalence” (p.375)

In the experiment described, there are two two-term contingencies - in one response R1 produces the reinforcer SR1 and in the other response R2 produces the same reinforcer. The baseline also contains four defined three-term contingencies with the discriminative stimuli B1, B2, D1 and D2. Each two-term contingency is controlled by two discriminative stimuli, both B1 and D1 independently control R1, and B2 and D2 independently control R2. Thus according to the definition above, B1 and D1 will be related by equivalence, as will B2 and D2. The AB conditional discrimination has already established equivalence relations between the stimuli A1 and B1 and between A2 and B2. Thus, if the A1B1 and B1D1 pairs are members of the same equivalence relation then A1 and D1 must also be related by equivalence, and if A2B2 and B2D2 are members of another equivalence relation, the A2 and D2 stimuli will again be related by equivalence. In this case, when shown the D stimuli, it would be possible for the subject to select the appropriate A stimulus without the differential responses R1 and R2. Therefore given the equivalence test, the D stimuli may be functioning as samples rather than the responses R1 and R2. Thus it may not be possible to arrange an equivalence test where the differential responses function as the conditional (sample) stimuli as the responses will always be under the control of other stimuli.

Given this difficulty, Sidman then suggested the problem could then be viewed in a different way. As the responses cannot be separated from their controlling stimuli, then why not include both in the equivalence relation. The same reasoning can then be applied to the three-term relation:

“If a defined response that is controlled by a conditional stimulus can be included in the equivalence relation, why not also the defined response that is controlled by a discriminative stimulus ?” (p. 378)

This is a reasonable proposition as, although the derived responses R1 and R2 cannot satisfactorily be separated from the stimuli, as Sidman points out, the only way the D stimuli could become equivalent to the A stimuli was via the different responses. So although it is difficult to isolate the responses, “only by including those responses as elements of the event pairs that define the equivalence relation could we predict the emergent DA conditional discriminations” (p.380).

One useful consequence of including differential responses as part of the equivalence relation is that it removes the need for response mediation theories of equivalence. If the defined relations are included as elements of the event pairs comprising an equivalence relation then the equivalence relations will include all the directly taught and emergent stimulus-stimulus, stimulus-response, response-stimulus and response-response pairs. Therefore, Sidman suggests, to predict the emergent relations all that is needed is the behavioural definition of equivalence.

Sidman also suggests that if responses are included in equivalence relations, then in the context of an equivalence relation there is no need for a distinction between stimuli and responses. Rather, an equivalence relation should be seen as being made up of pairs of events "with no restriction on the nature of the events that make up the pairs" (p. 384)

Analysis of the elements that make the equivalence relation in this way led Sidman to re-affirm his earlier suggestion that equivalence is a basic stimulus function. While there is some evidence that the pre-requisites for equivalence relations may be established at the level of the three-term unit, evaluation of the defining properties of reflexivity, symmetry and transitivity can only be done at the level of the four-term contingency. However, this is not particularly a problem as investigation at this level has shown just how useful the defining properties of equivalence are at predicting the development of equivalence relations. The experiments on responses and reinforcers have shown that all four elements of the four-term analytic must be included in the equivalence relation. The question then is, where does the unit come from; what is the basis of stimulus equivalence? Sidman believes that the establishment of equivalence relations is one of the outcomes of reinforcement contingencies. While reinforcement contingencies generate equivalence relations, reinforcement itself is a product of survival contingencies, thus supporting Sidman's earlier proposal that we form equivalence relations because we are "built that way". Thus, if equivalence is an outcome of reinforcement this confirms Sidman's original (1990) suggestion that equivalence is a primitive function.

While the evidence above documents how Sidman's belief that equivalence was a basic stimulus function was strengthened by further analysis of the behavioural units involved in equivalence relations, his other belief, that equivalence emerged only at the level of the four-term contingency, changed as a result of further research. In his 1994 book on equivalence he stated that he now believed there was

good evidence for the establishment of equivalence relations at the level of the three-term contingency. This requires considering behaviour at the level of simple rather than conditional discriminations.

As a result of re-evaluating research on equivalence relations and three-term contingencies Sidman (1994) published a retraction of conclusions in an earlier paper on this topic - Sidman, Wynne, Maguire and Barnes (1989). In that paper, Sidman et al argued that functional classes established by a series of simple discriminations (three-term contingencies) represented *different* behavioural processes to equivalence relations established by four-term conditional discriminations. Now however, Sidman believes that this conclusion is inappropriate and that, while the relations established in functional classes are defined and tested differently, they do imply equivalence relations in behaviour.

The issue of functional classes and equivalence relations arises mainly from consideration of the study by Vaughan (1988) investigating the possibility of equivalence relations with pigeon subjects (see chapter 3). The main point of theoretical debate centred on Vaughan's definition of equivalence. The behavioural definition of equivalence, proposed by Sidman and Tailby (1982), which requires the demonstration of the properties of reflexivity, symmetry and transitivity, was based on the mathematical formulation of equivalence. The demonstration of these properties requires the use of conditional discriminations (four-term contingencies). However, Vaughan pointed out that there is another way to view equivalence, also based on a mathematical definition, and that is as a "partition". Using a partition as the definition of equivalence allows equivalence relations to be investigated using a series of simple discriminations (three-term contingencies) which then form "functional classes".

The principle of partition is that

"If R is an equivalence relation on a set S , then there exists a partition P of S such that [elements of the set] lie in the same class P if and only if aRb holds.

Conversely, if P is a partition of S , then the relation . . . is an equivalence relation" (Gellert et al, 1977, p.324, cited in Sidman 1994, p.417)

or as Sidman puts it "a partition implies an equivalence relation and an equivalence relation implies a partition" (Sidman, 1994, p.417).

Initially Sidman was not convinced that a partition implies equivalence in the same way that equivalence is shown in a conditional discrimination paradigm. However,

in his subsequent re-evaluation he states that not only does a partition definition of equivalence make intuitive sense, but also that a correspondence between the two proposed definitions of equivalence actually strengthens the behavioural definition of equivalence. The mathematical definition of an equivalence relation and a partition do integrate, so if the behavioural definitions did not integrate in the same way then the utility of the definitions would be weakened; what would be one process in mathematics would have to be dealt with separately behaviourally.

The partition definition of equivalence makes sense in that partition is another way of viewing *classification*. If a group of stimuli is partitioned/classified off from other stimuli, then there must be some basis for that separation. Therefore, the elements of the group will be equivalent with respect to the basis on which they were separated. For instance, if a group of objects of differing shapes and sizes were classified according to colour, one group could be seen as equivalent because they were all red, another group would be equivalent on the basis of being green etc. Alternatively stimuli could be classified on the basis of their function. This is what happened in the study by Vaughan (1988). The stimuli were arbitrarily assigned to two groups. These groups formed functional classes depending on whether they did or did not produce reinforcement at any given time.

Sidman (1989)

The experiment by Sidman et al (1989) was based on the procedure developed by Vaughan (1988) which tested equivalence using a partition. In considering whether a partition (which established functional classes) was behaviourally the same as an equivalence relation (based on the defining properties of equivalence), one question was whether members of equivalence classes would also form functional classes. Several studies had already shown that this was true (e.g. Lazar 1977, Lazar and Kotlarchyk 1986, Mackay 1985, Wulfert and Hayes 1988). What Sidman et al were asking in this experiment was whether members of a functional class would show equivalence relations based on the properties of reflexivity, symmetry and transitivity.

This experiment was carried out with three subjects, one normally able adult, DJK, and two teenagers who were students at the New England Centre for Autism, PJV and JDB. Subject DJK was taught using upper and lower case Greek letters as stimuli. Subjects PJV and JDB were taught using numerals as stimuli. Neither of these subjects were familiar with the concepts "odd" or "even" and so the stimuli were assigned to classes on this basis. One test of the procedure then would be to

see if the subjects successfully assigned the stimuli used in training and testing to the appropriate classes at the end of the experiment.

The basic procedure was the same for all subjects. First, two functional classes were established using a series of simple discriminations. Two stimuli were presented to the subject; selection of one provided reinforcement (S+) and the other did not (S-). A series of simple discriminations were trained in this way to produce two functional classes; S+ and S-. When performance was stable on this task, the contingencies were reversed so that the class which had previously been associated with reinforcement, S+, now became S-, while the class which had functioned as S- now became S+. These reversals were continued until the subject made an error just at the beginning of each reversal. The subjects were then tested to see if these functional classes had also established conditional relations among the stimuli within each class; given a stimulus from one class as a sample, the question was whether the subject would reliably match that sample with a comparison stimulus from the same class rather than selecting the comparison stimulus from the other class.

If these conditional relations between functional class members were demonstrated, the subjects were taught conditional relations between some of the functional class members and some new stimuli. The subjects were then tested to see if conditional relations had emerged between the new stimuli and the other members of the functional classes. If these relations had emerged it would suggest the establishment of equivalence relations. In the final test, the simple discrimination procedure was used again. This time the new stimuli were included in the stimulus pairs to see if they had become members of the same functional classes as the stimuli to which they had been shown to be related by equivalence.

All subjects demonstrated the emergence of functional classes following functional discrimination training. After each reversal of contingencies the subjects switched their selections to stimuli from the new S+ class, with only an error on the first trial. This demonstrated that the subjects had formed functional classes as reversing the contingencies for one pair of stimuli was sufficient to change responding to all the other pairs.

The next test was to see if the subjects would display conditional relations between the stimuli within each functional class. Before it was possible to test this the subject were given a baseline of reinforced identity-matching trials to ensure that

they would respond to the probe trials as conditional rather than simple discriminations. In the conditional discrimination probes only two stimuli from each class were used, leaving the others free for later equivalence tests. All subjects demonstrated appropriate conditional discriminations among the stimuli, although with varying degrees of ease.

However, this is not a definitive test for equivalence as the stimuli had already been established as members of the same functional class; therefore appropriate matching relations were likely to be demonstrated on these tests. In the next test then, equivalence relations were tested by establishing conditional relations where the samples were new stimuli and the comparisons were members of the established functional classes. The subjects were then tested for the emergence of equivalence relations where the stimuli were different members of the functional classes and the comparisons were the new stimuli. Two of the subjects, DJK and PJV, clearly demonstrated equivalence relations, responding at a high level of accuracy on probe trials. The third subject, JDB, did not demonstrate these equivalence relations. On several occasions performance on the baseline relations deteriorated and performance on the probe trials was at chance level. Following retraining and verbal instructions the baseline was reinstated and performance remained stable. However, even when baseline performance was stable, performance on the probe trials was around chance level, so his failure to demonstrate equivalence relations did not appear to be due to a breakdown of the functional classes. It would seem that although JDB formed and maintained the functional classes, and learned conditional relations between some class members and new stimuli, the new stimuli did not enter the established classes and equivalence relations could not be demonstrated. One explanation suggested for this was a breakdown of conditional relations between the members of the functional classes. Subject JDB was retested for conditional relations between the stimuli within each stimulus class, relations which he had previously demonstrated. However, when retested, JDB failed to demonstrate appropriate conditional discrimination relations.

The two subjects who had demonstrated appropriate equivalence relations were then given a slightly more stringent test for equivalence. In this test further new stimuli were related to the new stimuli from the previous equivalence test (former new stimuli). The subjects were given probes assessing relations between original functional class members and the most recently introduced stimuli. This test was more stringent as it meant that the relations between the former new stimuli and the

original class members, and between the new stimuli and the former new stimuli must all have been equivalence relations to allow accurate performance on probes between the newest stimuli and the original class members. Both subjects DJK and PJV demonstrated appropriate equivalence relations.

The final test assessed whether the new stimuli which had been related to the original functional class members by conditional discriminations had actually become members of the functional classes. The new stimuli were thus included in simple discrimination tests to see if subjects would respond to them in the same way they responded to the stimuli that comprised the original functional classes. All responses consistent with the experimentally defined contingencies were reinforced. Both subjects responded appropriately on these trials, indicating that the new stimuli had indeed become members of the functional classes. Therefore, the conditional discrimination training had transferred functional class membership to the new stimuli.

In the conclusions of this paper, Sidman et al stated that the failure of subject JDB to demonstrate equivalence relations showed that partitioning a set of stimuli into functional classes is *not* the same behavioural process as equivalence relations established by conditional discriminations. They felt that it was not surprising that they did not represent the same process as the behavioural definitions and behavioural tests differed so substantially. However, in the light of further research, Sidman (1994) retracted these conclusions. Sidman now believed that JDB's failure to demonstrate equivalence was a result of testing without reinforcement. This was not immediately obvious in that, even when performance on the baseline relations was stable, JDB still failed to demonstrate equivalence. It now seems likely that on the baseline relations, responding in the expected manner probably provided reinforcement in itself but this reinforcement was not available on test items. Without any indication of whether he was meeting expectations or not, JDB showed the usual effects of extinction. If this is accepted, the other two subjects' data provides support for the notion that the behavioural definition of a partition does in fact imply an equivalence class.

As the partition and the equivalence relation do not appear to be behaviourally distinct, the fact that their definitions and tests are so different becomes even more significant. As Sidman (1994) puts it, "a seemingly unlikely empirical confirmation of a relation marks a greater scientific advance than does a highly predictable confirmation" (p.421). As described previously, in mathematics a partition implies

an equivalence relation and an equivalence relation implies a partition. Now, Sidman believes that behaviourally a partition and an equivalence relation are parts of the same process, and this congruence between the behavioural and mathematical definitions of equivalence demonstrates the utility of the original mathematically derived definition of equivalence.

The evidence that the equivalence relation must include reinforcers and defined responses, and that functional classes establish equivalence relations and vice versa, led Sidman (1994) to conclude that:

"(a) Two or more discriminative stimuli that control the same two-term contingency will be related by equivalence; (b) discriminative stimuli that control different responses, even though correlated with the same reinforcer will partition themselves into different equivalence classes; (c) discriminative stimuli that are correlated with different reinforcers, even though controlling the same response, will partition themselves into different equivalence classes; and (d) equivalence relations can emerge from these three-term units" (p.416)

There seems to be good evidence then that equivalence relations can, and do, emerge as a result of simple discriminations; at the level of the three-term contingency. One problem with this suggestion though, is that much of the research demonstrating this has included four-term contingencies in the training and test paradigms. While the results from these experiments can be predicted and accounted for at the level of the three term unit, it would still be preferable to have a demonstration of equivalence based solely on three-term units. Sidman (1994) reports that experiments by Kawashima (1993) and Manabe and Kawashima (1993) provide good evidence that experiments of this sort will indeed demonstrate equivalence based on three-term contingencies. However, these results have yet to be published in full and thus are not available for general scrutiny. Sidman himself has also described an experiment based solely on three-term units that might conclusively demonstrate equivalence relations. However this experiment has not yet been carried out, although results from previous studies strongly suggest that it would indeed result in equivalence relations.

Sidman's theory of equivalence then is that equivalence is a basic stimulus function. It cannot be derived from any more basic functions, although it can be described at the level of the three-term unit of analysis, and is a result of reinforcement.

The Relational Frame Account of Equivalence

A third explanation for the emergence of equivalence relations has been put forward by Steven Hayes (1991). Hayes accounts for equivalence in terms of arbitrarily applicable relational responding, or relational frames.

Hayes (1991)

Hayes (1991) stated that there were several good reasons for interest in the phenomenon of stimulus equivalence. One reason is that there seems to be a correspondence between stimulus equivalence and various language phenomena. Dugdale and Lowe (1990) and Horne and Lowe (in press) have argued that language results in equivalence, while Sidman (1990, 1992, 1994) suggests that equivalence may underlie language abilities.

A second reason for interest in equivalence is that Hayes suggests that equivalence is unexpected; "it would not be readily predicted from a three-term contingency formulation" (p.20). For instance, in a conditional discrimination the subject learns that there is a greater probability of reinforcement for selecting stimulus B1 in the presence of A1 than in the presence of A2. However, as Hayes points out, this does not necessitate there being a greater probability of reinforcement for selecting A1 in the presence of B1 than in its absence. "In the natural environment the contingencies supporting conditional discriminations rarely seem to be symmetrically arranged in this sense" (p.20). The example Hayes gives is that, in the presence of a lion, a primate may learn to approach the cover of a thicket rather than open savannah. However, the reversal of this relation "Given thicket, approach lion" would not have adaptive value.

While symmetrical contingencies may not occur frequently in nature, human subjects do readily display equivalence relations when tested, and stimulus equivalence requires the demonstration of symmetry. Hayes suggests that this occurs because stimulus equivalence is just one example of a specific type of behaviour-environment interaction, arbitrarily applicable relational responding, and is the result of "prolonged exposure to the contingencies of reinforcement operating in the verbal community" (Barnes, 1994, p.95).

Arbitrarily applicable relations are relations which are not based on formal characteristics of the stimuli, but rather on relations based on additional contextual cues or social convention. Relations which are based on formal characteristics of

the stimuli (non-arbitrary relations) are relations such as "bigger than", "darker than" etc. Arbitrarily applicable relations are not based on these formal characteristics of the stimuli but are instead specified by the context. One example of an arbitrarily applicable relation is the relation between an object and its name. The name "ball" is not related to a ball by virtue of the ball's physical form but rather by the conventions of the verbal community.

Just as Sidman and Tailby's (1982) definition of stimulus equivalence required the demonstration of the properties of reflexivity, symmetry and transitivity, Hayes too describes certain types of stimulus relation that will be shown in arbitrarily applicable relations. These relations are mutual entailment, combinatorial mutual entailment, and transfer of function. The definition of mutual entailment is that "a relation between two events involves responding to the one event in terms of the other and vice versa". (p.22). For instance, if A is better than B, then B is worse than A. Symmetry would be one example of a relation of mutual entailment, where the relation is one of co-ordination or "sameness".

Combinatorial mutual entailment requires that if there are relations between A and B, and between B and C, in a given context, the same sort of relation must be entailed between A and C. This relation is more than just a simple expansion of mutual entailment. "In mutual entailment the entailed aspect of the relation is always specified to the same degree as the specified aspect" (p.23). In other words, if A is better than B, then B will be worse than A to the same degree. However, in combinatorial entailment, while the relations between A and B and between B and C must both be specified, these may not in themselves be sufficient to specify the relation between A and C. For instance, A may be faster than B, and B may be larger than C, but this does not specify the relation between A and C. However, a relation between A and C *is* entailed. Transitivity and equivalence relations are both special examples of combinatorial mutual entailment.

The third characteristic Hayes specifies is that of transfer of functions. Hayes' rationale for this characteristic is that "relations between stimulus events would be of little importance to psychologists if the functions of these events could not in themselves be moderated by these relations" (p.23). For instance, relations might be established where A is smaller than B, and B is smaller than C. The properties of mutual entailment and combinatorial entailment specify the relations between A and C. Suppose then that stimulus A has a specific reinforcement function, relevant to the size of the stimulus. Transfer of function would then predict that stimulus B

and C would have ordinarily more reinforcement value than stimulus A. Therefore a subject accustomed to being rewarded with stimulus A would be expected to work harder when promised a reward of stimulus C, despite not having direct experience of reward with that stimulus. Similarly, an aversive function may transfer to a new stimulus without the subject directly experiencing the aversive consequences.

Hayes suggests that relations of arbitrarily applicable responding will display these properties, and that equivalence is just one example of these kinds of relations. To simplify discussion, Hayes has introduced the term "relational frame" to describe this kind of behaviour. The term "frame" is used to emphasise that this responding is not based on the specific stimuli involved; any kind of stimuli can be placed in the "frame". Instead, in the appropriate context arbitrarily applicable relational responding can be brought to bear on any set of stimuli. Thus a relational frame can be defined as:

"(a) a form of responding that manifests the contextually controlled properties of mutual entailment, combinatorial entailment, and transfer of function, (b) is produced by a history of relational responding appropriate to the contextual cues involved, and (c) is not produced by explicitly trained nonrelational responding with regard to the specific stimuli involved, nor by their physical properties alone" (Barnes, 1994, p.99-100).

As stated before, stimulus equivalence can be seen as just one example of a relational frame or arbitrarily applicable relational responding. In this case the frame is one of co-ordination or "sameness". Other types of relational frame might be "opposition", "distinction" or "comparison". Because of the type of relation involved in these frames, different patterns of relations are established among the stimuli on which they act. For instance, if a relational frame acts on four stimuli, A, B, C and D, and the relations are known between A and B, A and C, and C and D, then it is possible to say what other relations are entailed between the stimuli, and in some cases, to specify the form of these relations. In a frame of co-ordination A is the same as B and C, and C is the same as D. Therefore, A will be the same as D, and B and C will be the same as each other, and also the same as D. This describes a typical equivalence relation (see Fig.5:6). In a frame of opposition, A is the opposite of B and the opposite of C, and C is the opposite of D. Mutual entailment will thus show that C and B are both the opposite of A, and D is the opposite of C. Combinatorial entailment shows that C and B must be the same (both are the opposite of A). A and D must be the same, as both are the opposite of B and C (see Fig.5:7). In both these frames, co-ordination and opposition, it is

possible to specify the relation between the stimuli. This may not be possible in all relational frames. For example, two other relational frames might be "distinction" and "comparison". In a relation of distinction the stated relations might be that A is different to C and different to D. While a relation is entailed between stimuli C and D it is not possible to specify this relationship; they might be the same as each other *or* different to each other. Similarly, in a frame of comparison, the stated relation might be that A is better than C and better than D. Again, although a relation is entailed between C and D, it is not possible to specify the value of C and D relative to each other.

Hayes also believes that relational classes can develop from a mixture of relational frames, and in this way entire relational classes can be related to other relational classes. "For example, if one equivalence class is the opposite of another equivalence class, then each member of the first class is the opposite of all members of the second and vice versa" (p.35).

Steele and Hayes (1991)

Evidence for a relational control theory of responding comes from a study by Steele and Hayes (1991) where subjects (normally able teenagers) given an appropriate pre-training history responded on arbitrary matching to sample tasks in ways that appear to fit with frames of co-ordination, distinction and opposition. Unfamiliar visual forms were used as contextual cues for the subjects to relate stimuli according to relations of "sameness" (e.g. large square with a large square), "opposition" (e.g. large square with a small square), or "distinction" (e.g. square with a cross). The subjects then received conditional discrimination training, each discrimination being trained in the presence of one of the three contextual cues. For instance, the subject might be given the following training trials :

O A1 B1 B2

O A1 C1 C2

where O is the contextual cue for opposite, A1 is the sample and the B and C stimuli are the comparisons. The subjects were reinforced for choosing the underlined stimuli.

A test trial might then be:

O B2 C1 C2.

If the subjects responded according to equivalence then they should select C2 as both C2 and B2 are associated with A1. Or, if responding was controlled by the O stimulus they should select C2 as selecting C2 had already been reinforced in the presence of the O stimulus. However, if the subjects were responding according to a relational frame of opposition, they should select C1. This is because C2 and B2 are both the opposite of A1, and therefore B2 and C2 are the same. Thus the subject should select C1 as the opposite of B2. This is what was found, providing evidence that the subjects were responding according to relational frame theory. Hayes (1991) sees this as evidence to support the proposition that stimulus equivalence is one example of a class of phenomena that can be derived based on relational frames. In the case of stimulus equivalence, responding is based on the frame of co-ordination.

Thus, relational frames are examples of arbitrarily applicable relational responding. This is responding which is not based on the form of the relatae but rather brought to bear by other events, such as language or social convention.

Barnes (1994)

A paper by Dermot Barnes (1994) reviewed Hayes' concept of relational frame, and considered some of the differences between this and Sidman's account of equivalence. Several of his comments are also applicable to the naming account of equivalence by Dugdale and Lowe (1990) and Horne and Lowe (in press).

Sidman views equivalence as a basic stimulus function which precedes language, although it requires some exposure to verbal contingencies before it can be measured procedurally. Relational frame sees equivalence and language as representing the same derived processes of arbitrarily applicable relational responding. A naming account sees equivalence as emerging as a result of verbal behaviour. Barnes has suggested that, at least between relational frame and Sidman's theory, these distinctions may effectively be unimportant. As both theories suggest some exposure to verbal contingencies is necessary before equivalence can be tested this distinction may be unimportant in terms of empirically examining the differences in the theories, a suggestion which would also seem applicable to the naming account of equivalence.

Another difference would seem to be that relational frame considers equivalence as one of a number of derived relations, while Sidman's account regards equivalence as the most important or fundamental relation. However, again Barnes argues that

the differences may be more apparent than real. While relational frame theory does not treat equivalence as unique it does regard it as possibly the most important relational frame as it is central to referential relations in natural language. Sidman also sees equivalence as important in the concept of reference. Barnes suggests then that:

"Both accounts consider equivalence (a) as one of a number of derived relations, (b) as the most important or fundamental relation, and (c) to be critical for the development of referential relations in natural language". (Barnes, 1994, p.103)

Similarly, while Sidman's procedural definition of equivalence does not incorporate contextual control in the same way as Hayes' account, contextual control is fundamental to Sidman's account of equivalence.

Thus, Barnes shows that while there seem to be substantial conceptual differences between two of the accounts of equivalence, in several cases these differences are either not as substantial as they would seem, or difficult to test empirically. Barnes suggests that to differentiate between the accounts it is necessary to make a detailed examination of how each account would deal with certain anomalies found in equivalence testing. For instance, some studies have found superior equivalence performance following a multiple-sample single-comparison (many-to-one) training procedure than following a single-sample multiple-comparison (one-to-many) training procedure. Barnes suggests that relational frame theory may be better equipped to deal with this finding. Thus, the most profitable way of investigating the basis of equivalence is not to compare the theories on theoretical grounds, but rather to consider how well each theory copes with findings from ongoing equivalence research.

Figure 5:1 The two-term (Reinforcement) contingency.
R = Response; C = Consequence.
From Sidman (1986)

R1 (PRESS) \longrightarrow C1 (COIN)

R2 (OTHER) \nrightarrow C1 (COIN)

Figure 5:2 The three-term (Discrimination) contingency.
 S = Stimulus; R = Response; C = Consequence.
 From Sidman (1986)



Figure 5:3 The four-term (Conditional Discrimination) contingency
 S = Stimulus; R = Response; C = Consequence.
 From Sidman (1986)

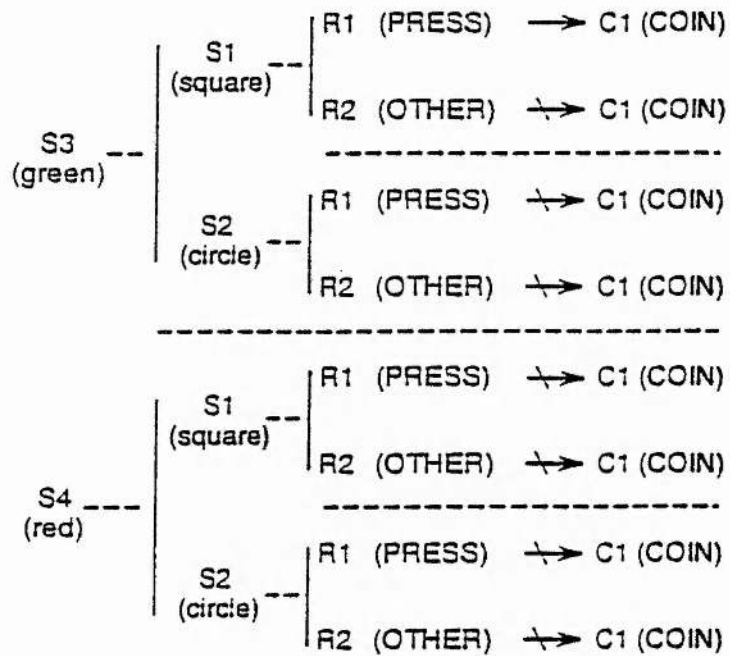


Figure 5:4 A balanced four-term (Conditional Discrimination) contingency.
 S = Stimulus; R = Response; C = Consequence.
 From Sidman (1986)

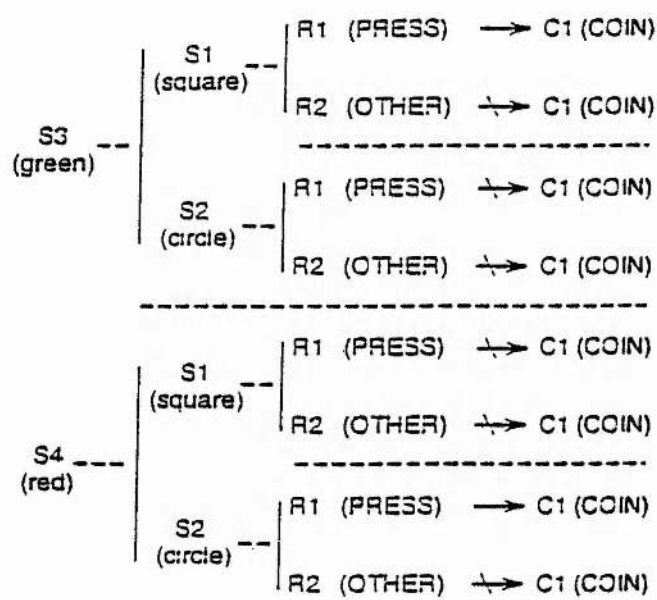


Figure 5:5 A proposed test for the inclusion of defined responses in the equivalence relation.
From Sidman (1994)

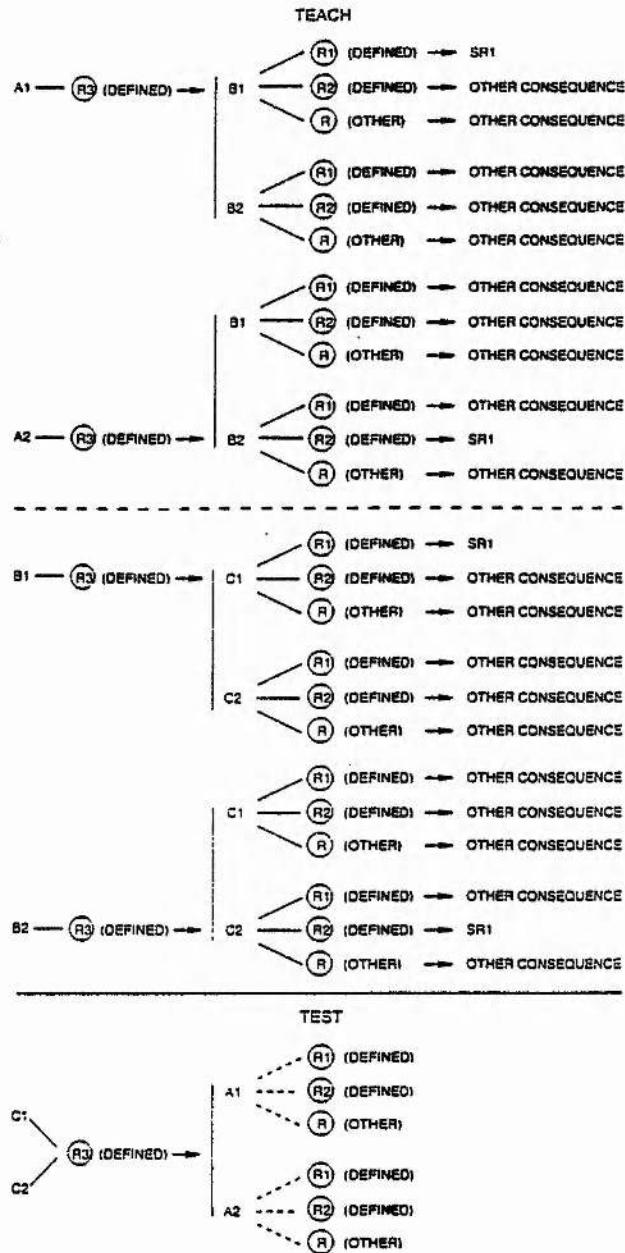


Figure 5:6 Archetypal Coordination Network
From Hayes (1991)

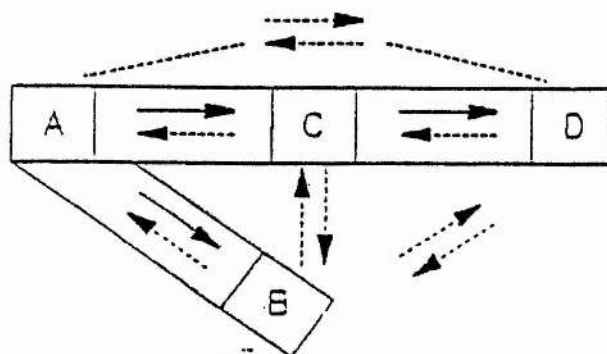
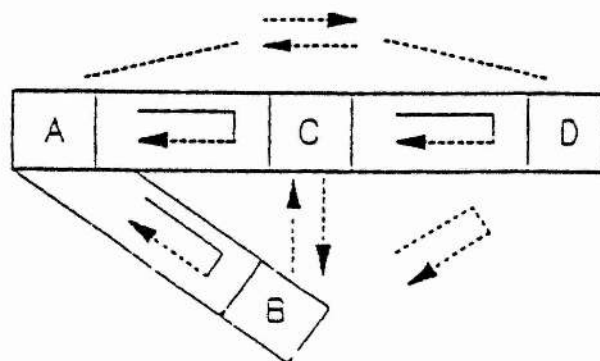


Figure 5:7 Archetypal Opposition Network
From Hayes (1991)



CHAPTER 6

Experiment 1:

Investigation of the Role of Naming in Equivalence Performance

Investigations of performance on stimulus equivalence tasks has suggested that there may be some connection between language and equivalence. After all, one of the very first experiments into equivalence-like phenomena (before the formal definition of equivalence relations by Sidman and Tailby, 1982) was by Sidman (1971), which documented the emergence of 40 new untaught relations after a subject with learning disabilities was taught to match visual words and visual pictures to 20 dictated names. Even at this early stage in investigation into stimulus equivalence it seemed possible that the behaviour being observed might be similar to that in natural language acquisition.

While the defining feature of stimulus equivalence is the emergence of new untaught behaviours from explicitly established relations, it has been noted repeatedly how closely these equivalence relations seem to correspond to language phenomena. The emergent relations between spoken or printed words and pictures or objects seems to relate to the natural learning of names for objects. Equivalence relations seem to closely match the relations implicated when we say that a symbol "stands for" an object or that a name "means" that object.

What is not clear is the exact nature of the relationship between stimulus equivalence and language. Sidman (1990; 1992; 1994) has proposed that equivalence is a basic stimulus function which cannot be derived from any more basic process. He suggests that language and verbal behaviour develop from the more basic equivalence relations. Other researchers (e.g. Dugdale and Lowe, 1990) have argued that in fact language underlies the development of stimulus equivalence and that the emergence of equivalence relations is a direct result of verbal behaviour. Most specifically, Dugdale and Lowe (1990) suggested that equivalence was a direct result of the establishment of bi-directional relations between stimuli and names.

One of the strongest sources of support for a verbal behaviour interpretation of equivalence has been the difficulty in demonstrating emergent equivalence relations with any non-human subject. Dugdale and Lowe had themselves tried to establish symmetrical responding with three chimpanzee subjects. All three

chimpanzees had participated in an ape-language training programme, and had an extensive pre-history of training and testing on abstract relations. Despite this, the chimpanzees gave no indication of having derived symmetric BA responding as a result of being directly trained on an AB conditional discrimination. Other attempts to demonstrate equivalence relations with non-humans have shown a similar lack of success, although there is now some evidence that Schusterman and Kastak (1993) may have demonstrated equivalence with a sea-lion subject. However, the difficulty in demonstrating equivalence relations with any non-human subjects, compared to the frequency with which human subjects, even those with quite severe learning disabilities, display equivalence responding, has been seen as evidence for the importance of language in the formation of equivalence relations.

Another apparent source of support for the importance of naming in equivalence has been the finding that subjects who initially fail to demonstrate equivalence relations will often demonstrate these relations after being taught to name the stimuli used in the baseline conditional discriminations. (e.g. Lowe and Beasty, 1987; Dugdale and Lowe, 1990). However, several of these studies reported a correlation between language/naming and equivalence performance rather than demonstrating a causal relationship. One study which tried to examine the precise nature of the relationship between naming and equivalence performance was Eikeseth and Smith (1992). This study used a five-stage procedure to examine equivalence performance, with and without programmed naming, with four autistic children. In phase 1, the subjects were taught conditional discriminations without programmed naming. If the subject failed to demonstrate equivalence, in phase 2 they were taught to name the stimuli while responding on the conditional discriminations, and tested once more for equivalence responding. In phase 3, the subjects were taught names for a new set of stimuli without establishing conditional relations between the stimuli. They were then tested for the development of conditional relations, and if necessary these relations were established before testing for equivalence. This assessed if naming alone was sufficient to establish equivalence responding. In phase 4, two new stimuli (without programmed naming) were related to the stimuli used in phase 3. The subjects were then tested to see if; (a) conditional relations would develop between the new unnamed stimuli and the named stimuli from phase 3, and (b) would conditional relations develop between the new unnamed stimuli. Phase 5 was the same as phase 1. Using a new set of stimuli conditional relations were trained, with no programmed naming, and the subjects were tested for the emergence of equivalence relations (see chapter 5 for a fuller description and discussion).

One subject seemed to show a causal relation between naming and equivalence relations, consistently demonstrating equivalence relations when naming was programmed and failing to display equivalence when it was not. The other subjects' performances were not as consistent in this respect, although they did appear to be more likely to display untrained conditional relations when naming was programmed compared to when it was not programmed. However, the results of this study were not conclusive and it is not clear whether naming causes equivalence as Dugdale and Lowe (among others) have suggested or whether equivalence results in naming as Sidman suggests.

The aim of this experiment is similar to that of Eikeseth & Smith (1992): to systematically investigate the role of naming in Stimulus Equivalence. This was to be done in two ways: (i) By measuring the amount of spontaneous naming behaviour shown, and (ii) If the subjects failed to demonstrate stimulus equivalence they were taught to name the stimuli and then retested for the emergence of stimulus equivalence.

If equivalence is a fundamental stimulus function as Sidman suggests, then subjects who show emergent stimulus equivalence relations need not necessarily display evidence of naming. If Dugdale and Lowe are correct, and the demonstration of equivalence requires naming, then subjects displaying stimulus equivalence should show good evidence of naming behaviour. Further, teaching subjects to name the stimuli should be sufficient to produce stimulus equivalence in subjects who have previously failed equivalence tests.

While this experiment has similarities to that by Eikeseth & Smith, there are certain important differences.

The subjects in the study by Eikeseth and Smith were high functioning autistic children. The authors suggested that as these subjects often display unusual learning characteristics there may be difficulties in generalising results from these subjects to a wider population. This study then used adults with learning disabilities as subjects. Adults with learning disabilities appear less likely to display emergent behaviours, such as stimulus equivalence, than normally able adults and therefore examination of their performance on tasks like these may give some idea of the processes underlying equivalence. Moreover, any effects on performance are likely to be due to experimental manipulation, rather than

developmental factors as might be the case with normally developing young children.

Eikeseth and Smith had used a trial and error procedure to train the baseline conditional discriminations, but found that in the initial stages of the study, hundreds of trials were required before the subjects mastered the discrimination. One subject required 1,546 trials to complete the AB and AC conditional discriminations in phase 1. Eikeseth and Smith had therefore suggested that a more efficient training procedure should be used in any subsequent studies. Research in transitive inference studies has shown that success on tests for transitive inference is in direct relation to retention of the relations trained (Bryant & Trabasso, 1971). It is likely that stability of the training relations is equally important in stimulus equivalence. For this reason an errorless learning shaping procedure was used to train the baseline conditional discriminations in the present study. Ensuring stability of the training relations means that any failure to demonstrate stimulus equivalence is a genuine failure and not a result of incomplete training.

Method

Subjects

A total of seven adults with learning disabilities took part in a number of the experiments recorded in this thesis, several taking part in more than one experiment. Six of these subjects were recruited from a residential hostel in St. Andrews. Four of the subjects lived in the hostel itself, and two others lived semi-independently in a house attached to the hostel. Data was collected for these subjects, to try and get some idea of their cognitive and living abilities outside the experimental test situation. This data was collected from a number of independent tests, such as the BURT word reading test and the British Picture Vocabulary Scale, and from the care staff in the hostel. This data is summarised in Table 6:1.

The seventh subject, GT, had been recruited separately as he was already being tested on a number of independent clinical measures, and these had suggested that it might be interesting to examine GT's performance on a stimulus equivalence task. As a result GT was not tested on the same independent measures as the other subjects, but data is available on his performance on the Wechsler Adult Intelligence Scale (WAIS), and the Wechsler Memory Scale - Revised (WMS-R;

see Table 6:2). GT lived fairly independently in a house in a town near St. Andrews. On the WAIS GT's verbal IQ was found to be 92, which falls in the normal range, while his performance IQ was lower - 74. On the WMS-R, GT scored in the bottom 10% of the population overall. On the paired associates task within the WMS-R, which may have particular relevance for equivalence relations, GT's verbal and visual scores were around chance level.

Independent Tests

The subjects were tested on the BURT Word Reading Test, the British Picture Vocabulary Scale (BPVS) - short form, the British Abilities Scales (BAS) - matrices, and the Test for Reception of Grammar (TROG). The BURT Word Reading Test was used to get some idea of whether the subjects were able to read. This might have been relevant as some of the stimuli were printed words that were assumed to be unfamiliar to the subjects. The British Picture Vocabulary Scale gives some idea of the subjects' ability to select the appropriate picture when given a dictated name. The sample words become increasingly difficult as testing continues. It was possible that this measure would be relevant to the subjects' naming abilities. The matrices test from the British abilities Scales was used to get some measure of the subjects' visual reasoning abilities. The Test for Reception of Grammar gives some measure of each subjects' language comprehension, assessing increasingly complex aspects of grammar. For instance, the first three blocks test comprehension of nouns, verbs and adjectives respectively., while later blocks test items such as negatives, "X but not Y", and "not only Z but also Y".

Assessment by Care Staff

The care staff gave an assessment of the amount of support each subject required in day to day living, rating each subject as high dependency, moderate dependency or low dependency. They also gave an indication of whether they felt the subject could read, could write, or could copy their own name, or whether they made a mark as a signature.

Subjects Taking Part in this Experiment

Four adults took part in this experiment. There were subjects, JD, TD, NL (see Table 6:1), and subject, GT (see Table 6:2).

Stimuli

The stimuli used were eight Greek letters. These were used as it was unlikely that the subjects would be familiar with the stimuli, or would have any knowledge of

the relations between the stimuli, prior to training. The upper-case and lower-case characters and printed names of each letter were used. This gave eight three-member stimulus classes (see Table 6:3). All 24 stimuli were used in naming tests given before and after training. Two classes of stimuli at a time were used for training and testing. The stimuli were presented on photographic cards, approx. 13cms x 9cms, covered with clear plastic. Each card represented one stimulus. As an errorless training procedure was used, six cards were prepared for each individual stimulus, at different levels of intensity. These were prepared photographically at different exposures so that each stimulus ranged from pale grey on white at the lowest intensity, to full black on white at full intensity.

Design

The study was carried out in two phases, each preceded and followed by a general naming test.

Phase 1

In phase 1 the subjects were taught four conditional discriminations, A1-B1, A2-B2, A1-C1, A2-C2. They were then tested for the emergence of B-C and C-B relations which would be indicative of the formation of equivalence classes. This phase was designed to see if the subjects would display stimulus equivalence on the basis of the conditional discriminations trained alone, as no naming was programmed for this phase.

Phase 2

In phase 2, a new set of stimuli was used to teach another four conditional relations, A3-B3, A4-B4, A3-C3, A4-C4. This time the subjects were required to name the stimuli as they learned the conditional discriminations. The subjects were then tested on the B-C and C-B relations to see if naming facilitated the emergence of equivalence relations.

Procedure

Pre and post training naming tests

Before any training was given the subjects were tested for any evidence of spontaneous naming of the stimuli. The 24 (full intensity) stimulus cards (8 upper case, 8 lower case, 8 printed name) were presented one at a time to the subjects. The subjects were asked to say if they thought that stimulus had a name or "should be called something". They were also permitted to say that they did not think any particular stimulus had a name, or if they could not think of it. This naming test was repeated at the end of phase 1, and again at the end of phase 2 if the subject took part in that phase. This was to see if the subject showed any consistent use of

names from pre- to post- test or if the subject was more likely to name the stimuli following training with or without naming. There were no "correct" or "incorrect" names for the stimuli. The test was to check if the subjects would spontaneously name the stimuli, and would they use these names consistently.

Training without naming

An errorless learning shaping procedure was used to train the conditional discriminations. In this procedure, one salient aspect of the test setting is enhanced and used to establish appropriate responding to the correct stimulus. The environment is then shaped back to the full test situation, hopefully while maintaining correct responding. (Aeschelman & Higgins, 1982; Jones & Cullen, 1980; Luiselli & Donellon, 1980; Schilmoeller & Etzel, 1977). On each trial a sample stimulus was presented (e.g. A1), followed by two comparison stimuli (e.g. B1 and B2). On each trial the sample and correct comparison stimuli (S+) appeared at full intensity black on white. The incorrect comparison (S-) initially appeared as pale grey on white. As correct responding to S+ was established, the incorrect comparison slowly darkened in colour until it too was at full intensity black on white. There were six different levels of intensity. Three consecutive correct responses were required at each level before the intensity was increased. An incorrect response at any level resulted in the intensity of the S- stimulus immediately being reduced by one level until performance was stable once more. Once both the correct and incorrect comparison stimuli were at full intensity six consecutive correct responses were required before the subject was said to have reached criterion on that discrimination.

On all trials the sample stimulus was placed in front of the subject and the subject was required to look at the stimulus and touch it. The two comparison stimuli were then presented side by side underneath the sample stimulus and the subject was asked to point to the stimulus they thought "went with" the sample stimulus. Following a correct selection the subject was told "Yes, that's right, well done". Following an incorrect selection the subject was told "No, that's wrong, it's this one" and the experimenter pointed to the correct comparison. The left/right position of the comparison stimuli varied randomly from trial to trial.

Initially the subjects were taught A1-B1 and A2-B2 separately. Once the subject had reached criterion on both these tasks, they were combined to form the A-B mix. In this task the sample stimulus switched between A1 and A2 in a pseudo-random sequence, so that each stimulus appeared five times in a block of ten trials.

The position of the comparison stimuli varied randomly from trial to trial. All stimuli were presented at full intensity. No feedback was given on these trials. Criterion on this task was set at 90% correct responding. After the A-B relations were trained, the A1-C1 and A2-C2 relations were trained in the same way and the A-C mix was tested.

Testing

The same procedure was used for testing the A-B and A-C mixes as for the B-C and C-B equivalence relations. On these trials the sample stimuli varied in a pseudo random sequence so that each relation was tested equally often. The left/right position of the comparison stimuli varied randomly from trial to trial. Criterion on all tests was set at 90% correct responding. Before a test trial block, and before the A-B and A-C maintenance tests were given, the subject was told "This time I'm not going to tell you if you're right or wrong, but you've done very well, so just do your best".

Training with naming

If the subjects did not demonstrate emergent equivalence relations following training on phase 1 they progressed to phase 2 of the study. In phase 2, the subjects learned four more conditional relations with two new sets of stimuli, A3-B3, A4-B4, A3-C3, A4-C4. The basic training procedure was the same as in phase 1 but this time, when responding on the training trials, the subjects were required to label the stimuli. Thus, when training relation A3-B3, the subjects would be shown stimulus A3 - upper case Theta. The subject had to point to the sample and label it "theta". This would then produce the comparison stimuli B3 - lower case Theta and B4 - lower case lambda. For a response to be termed correct the subject had to point to B3 and label it "theta". If the subject selected the wrong comparison, or gave the wrong label, the response was scored as incorrect. Following an incorrect response the subject was told "No, this one [B3] goes with this one [A3], this one [B3] is called theta".

A similar procedure was used to teach the A4-B4, A3-C3 and A4-C4 relations. On test trials the subjects were not required to label the stimuli when responding, but were not prevented from doing so. Responses on test trials were scored as correct or incorrect based on the stimulus the subject pointed to, without reference to any name they may or may not have produced during responding.

Results

The results for each subject will be presented individually

Subject JD

Phase 1

Subject JD learned the four baseline conditional discriminations easily, reaching criterion on the first training session of each discrimination. Responding on these discriminations was between 95% and 100% correct (see Fig. 6:1). JD also achieved criterion on the A-B and A-C mixes the first time they were tested. JD responded at 100% correct on the B-C and C-B tests for equivalence. Thus JD clearly demonstrated emergent equivalence performance.

Naming

On the pre-training naming test JD gave names for 5/16 Greek letters. Three of these stimuli were subsequently used in training and testing. JD did not give names for any of the printed words. On the post-test naming check JD gave names for 14/16 Greek letters, although she had not seen 12 of these letters since the previous naming test. She also spelled out the letters of two of the printed words. Neither of these words had been used in training.

Subject TD

Phase 1

Subject TD learned the baseline A-B and A-C conditional discriminations easily, responding at between 96% and 100% correct (see Fig. 6:2). Although TD was tested twice on the A-B mix she failed to reach criterion, scoring 63% correct on the first test and 45% correct on the second. As it was not possible to test for equivalence without accurate performance on the A-B and A-C mixes, TD progressed to phase 2 of the study.

Phase 2

In phase 2 TD was able to learn the new baseline discriminations easily while labelling the stimuli correctly, responding at 100% correct on these trials (see Fig. 6:3). Initially TD's performance on the A-B mix was around chance level (67%; 40%). However when the A-C mix was tested, performance on this approached criterion (83%). The A-B mix was retested and performance on this eventually also approached criterion (50%, 88%). TD was then tested on the C-B equivalence test and demonstrated emergent equivalence relations, responding at 100% correct. TD terminated the test session immediately after the C-B test and so the B-C relations were not tested.

Naming

In the pre-training naming test TD produced names for 6/16 of the letters, all for lower case letters. She also tried to pronounce 5/8 names - e.g. "oba" for omega, "philip" for phi. Immediately after equivalence testing TD terminated the test session and it was not possible to give the post-test naming check.

Subject NL

Phase 1

Subject NL seemed to have more difficulty than the other subjects in learning the conditional discriminations. Her performance on the A-B relations was good, reaching criterion on both A1-B1 and A2-B2 on the first attempt (see Fig. 6:4). However when the A-B mix was tested, performance was at chance level. Performance on the A-C discriminations was poor. NL reached criterion on the first presentation of A1-C1, but subsequently failed to reach criterion on A2-C2. Performance was tested on several occasions and this pattern was repeated; accurate performance was shown on A1-B1, A2-B2 and A1-C1 training trials, but NL never reached criterion at a level higher than stage 3 out of 6 of the errorless training procedure on A2-C2.

The A-C discrimination required the subjects to match an upper case Greek letter sample to the printed name of that letter. As NL could not read it was possible she found it difficult to discriminate between the printed name comparisons, in this case sigma and omega. This seemed to be supported by her accurate performance on the A-B training trials. The original C1 and C2 printed word stimuli were therefore replaced with two different lower case Greek letters to form new A-C relations. NL reached criterion on both these A-C relations the first time they were presented (see Fig. 6:5). However NL never reached criterion on the A-B or A-C mixes. Performance on these relations remained at around chance level despite the fact that NL now had an extensive history of training on the A-B relations.

Rather than immediately progressing to phase 2 of the study, NL received testing on two of the components of stimulus equivalence; reflexivity and symmetry. This was to see if the conditional discrimination training had established any of the pre-requisites for equivalence. All the stimuli were tested for reflexivity and the B1-A1, B2-A2, C1-A1, C2-A2 symmetrical counterparts of the training relations.

Performance on these component relations was very inconsistent (see Fig. 6:6). NL appeared to display 100% correct reflexive matching relations on A2-A2, B1-B1 and C1-C1, but also displayed 0% correct performance on the reflexive relations A1-A1, B2-B2, C2-C2. This would suggest that NL was not in fact reliably demonstrating reflexive performance on any relations. On the symmetrical relations NL displayed apparently perfect symmetry on B2-A2, chance-level performance on B1-A1 and C2-A2 and responded at 10% correct on C1-A1, suggesting that she had formed inappropriate symmetrical relations. As NL's performance on the necessary components of stimulus equivalence was so inconsistent it would be safe to assume that she had not formed equivalence classes as the result of conditional discrimination training.

The training and testing of these relations had taken place on several occasions over a period of about 6 months. NL withdrew from the study at this point and did not receive phase 2 training.

Naming

On the pre-training naming test NL produced "names" for 22 of the 24 stimuli. However, only two of these were names. For 11 of the letters and all 8 printed names NL gave the stimulus a number, often using the same number for several different stimuli. On the post-test naming check NL gave names for 6 of the 24 stimuli. She did not use numbers to name any of the stimuli.

Subject GT

Phase 1

Subject GT was recruited separately to the other three subjects and initially received a slightly different training procedure. Results from some independent clinical tests had suggested that GT might not require an errorless learning procedure and so initially the A-B and A-C discriminations were taught using a trial and error procedure. In this procedure all stimuli appeared at full intensity black on white and GT was required to make five consecutive correct responses to reach criterion on a conditional discrimination. GT learned the A-B discriminations rapidly and displayed 100% correct responding on the A-B mix (see Fig. 6:7). GT was slower to reach criterion on the A-C discriminations and, although he did eventually display criterion responding on the A-C mix this performance was not as reliable as that on the A-B mix (63%).

It was decided to retrain the conditional discriminations with a new set of stimuli using the errorless learning procedure to try and ensure accurate baseline performance before GT was given equivalence testing. GT rapidly reached criterion on the A-B and A-C relations with the new stimuli (see Fig. 6:8). He passed the A-B mix the second time it was presented. However, once again performance on the A-C mix was close to chance level. GT was given the C-B equivalence test anyway, but as expected, performance was around chance level.

Phase 2

GT therefore progressed to phase 2 of the study, again using the errorless learning technique. GT rapidly learned the A-B and A-C conditional discriminations while naming the stimuli (see Fig 6:9). Initially performance on the A-B mix was at chance level, but when retested both the A-B and A-C mixes approached criterion. GT was given the B-C and C-B equivalence tests. The B-C relations were tested first and performance was poor (60% correct). The C-B relations were tested next and GT showed 100% correct performance.

Naming

GT was not given the pre-training naming test. On the post-training test GT gave a name or description to all 24 stimuli. He also tried to pronounce all the printed words, making errors on xi and phi.

Discussion

Of the four subjects tested, one subject (JD) displayed equivalence relations in phase 1 when no naming was programmed. The other three subjects failed to display equivalence in phase 1, and two of the three subjects (TD and GT) proceeded to phase 2 of the study where they were required to name the stimuli while learning the baseline conditional discriminations. Both these subjects subsequently displayed emergent equivalence relations. Interestingly GT did not display equivalence when tested first on the B-C relations but only on the subsequent test of the C-B relations. This would suggest that the equivalence relations emerged during the course of equivalence testing and were not fully established prior to testing.

The conclusions that can be drawn from the naming test data are quite limited. JD appeared more likely to name the stimuli at the end of the testing procedure than at the beginning. However, none of the names given at the beginning of testing

corresponded with any names given at the end. During both training and testing JD responded extremely quickly and was inclined to talk about events other than the test setting during testing. This would suggest that her attention was not focused on using a naming procedure to control her responding. In fact, none of the subjects were inclined to overtly name the stimuli except when explicitly required to do so by the experimenter.

For two subjects the naming procedure did appear to facilitate emergent equivalence performance. However, with subject GT this performance emerged during testing, and as the subjects were not required to name the stimuli during testing, it is not clear that the naming caused the emergence of equivalence relations. For both TD and GT it would be possible for the derived equivalence performances to be due to additional experience of the training and test procedures rather than due to the naming itself.

Subject NL, did not display equivalence relations in phase 1, despite an extensive history of training and testing. Interestingly, on the pre-training naming test, NL was inclined to try and give a name to every stimulus. In this case NL might have been inclined to spontaneously try and label the stimuli during discrimination training. However, she certainly did not do this overtly, and any covert naming does not appear to have facilitated accurate performance on tests of equivalence relations.

Sidman (1994) pointed out that, for a number of reasons, "sometimes . . . we see subjects - even highly intelligent subjects - failing completely to show evidence of equivalence relations after having mastered a baseline of four-term units" (p.407). Given that the same procedure can produce both successes and failures it is as important to consider why it might work as well as why it might fail.

When considering the results shown, it is important to consider both what is necessary to explain the results, and also what is sufficient. JD clearly demonstrated equivalence relations, but there is no clear evidence that she named the stimuli, let alone that her performance was dependent on naming. Certainly JD seemed more likely to name the stimuli in the post-test naming procedure than in the pre-test procedure. However, this was as true for stimuli not used in the experiment as for those used. This might as easily indicate a greater willingness to spontaneously name the stimuli as to indicate any naming process operating to produce equivalence. Also, JD did not produce names for any of the printed words.

If the stimuli used in training and testing had become equivalent through naming it would seem reasonable for these names to apply to all the stimuli used and there is no evidence that they did so. This would suggest that the increase in naming behaviour was coincidental rather than an indication of a causal process. While naming *might* provide an explanation of JD's performance, there is no clear reason to conclude that it does.

For GT and TD, the introduction of naming to the procedure did seem to facilitate the development of equivalence relations. However, this alone is not sufficient to identify naming as a necessary process for the development of equivalence relations. None of the subjects were inclined to overtly name the stimuli except when required to do so. Certainly none of them ever produced consistent group names for the two classes of stimuli, either during equivalence training and testing or during the naming checks. Thus, while some naming behaviour may have occurred, there is not sufficient evidence to conclude it was necessary or even important. Similar results were obtained by Eikeseth and Smith (1992). One subject did demonstrate equivalence relations while naming was programmed and failed to display equivalence when it was not. The other subjects did not show as clear a pattern although, overall, they were more likely to demonstrate equivalence when naming was programmed than when it was not. However, this general pattern is not sufficient to demonstrate naming is necessary for equivalence compared to possibly facilitating equivalence in some way.

It is worth noting that the evidence from the naming tests is inconclusive and it is possible that these tests would not constitute a sufficiently sensitive test in any case. Dugdale and Lowe (1990) suggested that pre- and post- test naming checks may not be sensitive to naming that the subjects naturally use when responding on equivalence tests. They suggested that experiments in their laboratory by Hird (1989) showed that the verbalisations made spontaneously during testing were quite often different to the names given during subsequent naming tests. Often subjects would give no evidence of using any names when tested with a post-test naming check. However, as none of the subjects in the present study showed any inclination to overtly name the stimuli, or verbalise about the stimuli in any way during training and testing, it is not clear that recording spontaneous verbalisations would have been informative in this case. It is possible that a more sensitive measure of naming behaviour might provide more useful information.

Lloyd Morgan's view that "In no case may we interpret an action as the outcome of the exercise of a higher psychical faculty, if it can be interpreted as the outcome of one which stands lower in the psychological scale" (Lloyd Morgan, 1909, p.56) seems relevant to the interpretation of these results. Naming behaviour might explain the emergence of equivalence, but there is no clear reason to assume that it does. In that case, is there another "lower" process which could explain the results?

Sidman has suggested that equivalence is a fundamental stimulus function - a product of survival contingencies; reinforcement. In that case, the demonstration of stimulus equivalence would be a product of reinforcement contingencies. Would this be sufficient to explain the results obtained?

The training in this experiment was aimed at producing stable performance on the A-B and A-C conditional discriminations. To achieve this the subjects received reinforcement for selecting, for example, B1 in the presence of A1, and B2 in the presence of A2. When reinforcement is contingent on a particular behaviour it increases the likelihood that the subject will show that behaviour again. Therefore, reinforcing the subjects on the baseline training relations was sufficient to produce stable performance on these relations. According to Sidman, these reinforcement contingencies are sufficient to also produce equivalence relations as stimulus equivalence, like reinforcement, is a fundamental stimulus function. "We form equivalence relations because we are built that way" (Sidman, 1994, p.389) and "The capacity for demonstrating equivalence relations can be regarded as a most useful gift from our inheritance" (Sidman, 1994, p.391).

The reinforcement contingencies in training establish the conditional discriminations. Research has shown that functions acquired by one element of an analytic unit are transferred to other elements of that unit. "Thus, if a conditional relation is also an equivalence relation and the sample stimulus in the equivalence class joins a syntactic class . . . the comparison that is related to that sample will also join the syntactic class" (Sidman, 1994, p.392). In this case the contingencies arranged in training would be sufficient to create classes of equivalent stimuli.

If you accept Sidman's proposal of equivalence as a fundamental stimulus function, then the reinforcement contingencies arranged during training would be sufficient to explain the demonstration of equivalence in this experiment. This would be true both for JD and for subjects GT and TD. There would be no need to postulate naming as the cause of equivalence.

Why then did naming appear to facilitate the emergence of equivalence for GT and TD, and why did NL fail to show equivalence? Sidman has suggested that while language may not be necessary for equivalence it may help facilitate it. For instance, labelling the stimuli requires differential responding to the samples and the comparisons. In this way it supports the establishment of the samples as a discriminative stimulus. This enhances the process established by the reinforcement contingencies. In this way language may facilitate equivalence class development.

This may also explain why NL did not show any of the components of equivalence. For instance, even though NL learned the individual A-B relations, she did not reach criterion on the A-B mix. This would suggest that training had not securely established the samples as discriminative stimuli. This would then make it impossible for her to derive discriminative stimulus classes. If naming had been programmed it might have facilitated this process.

It is possible then to explain the results obtained in this experiment in terms of equivalence as a fundamental stimulus function. Naming might *help* establish discriminative stimulus classes but this need not imply a causal role for naming in the development of equivalence relations.

However, it should be remembered that the naming tests may not have been sufficiently sensitive to identify all naming behaviour and so it would be unwise to dismiss the possible role of naming in equivalence relations.

One further subject, AS, underwent a similar training and test procedure. This subject also had learning disabilities but had received teaching in ancient Greek over a period of several years. This gave him an unusual history of familiarity with the stimuli used for training and testing, and might have had implications for his performance on the equivalence test. AS seemed most likely to be able to name the stimuli spontaneously, and might in fact already "know" the stimulus classes to be derived. Would he then demonstrate appropriate performance on the stimulus equivalence tasks? If so, would he show any evidence of naming the stimuli/ If AS did show appropriate performance on the initial equivalence test, he would then receive training and testing with a novel set of stimuli (not Greek letters). Would he be able to show appropriate equivalence performance with the new stimuli?

Whatever his performance would he show any evidence of naming, and if so, would it show any signs of influencing his performance?

AS received an extensive programme of training on the conditional discriminations used in this experiment, and testing for any evidence of naming behaviour. The details of this training and testing are contained in Appendix I.

It proved difficult to establish stable A-B performance. Even when stable performance was achieved on both the A-B and B-C relations, performance on both the equivalence test, and any components tested, was around chance level. AS's performance on the naming tests was similarly erratic. He gave no evidence of naming the stimuli consistently and if this happened covertly it did not appear to facilitate performance during training or testing.

Table 6:1 Independent assessment of subjects with learning disabilities taking part in experiments 1, 2, 4 and 5

	Subject	AB	JD	TD	NL	MM	DP
	Age	35	70	34	60	49	39
	Sex	male	female	female	female	male	male
staff assessment	Dependency	moderate	moderate	low	moderate	low	low
	read/ write	mark	good reader	good reader	copy /mark	good reader	good reader
independent tests	BURT word reading test	0	0	43	0	31	0
	BPVS (short form)	6	17	17	10	19	14
	BAS (matrices)	0	2	3	0	2	2
	TROG	2 blocks	7 blocks	7 blocks	2 blocks	10 blocks	4 blocks

Table 6:2 Subject GT: Independent assessments of ability

					WAIS		WMS-R	
Subject	Age	Sex	Dependency	Read/ write	Verbal IQ	Performance IQ	overall	verbal and visual score
GT	65	male	Low (assumed)	good reader	92	74	bottom 10% of population	chance level

Table 6:3 Training and Test Stimuli

Stimulus class	A	B	C
1	Γ	γ	gamma
2	Δ	δ	delta
3	Θ	θ	theta
4	Λ	λ	lambda
5	Ξ	ξ	xi
6	Φ	ϕ	phi
7	Σ	σ	sigma
8	Ω	ω	omega

Fig 6:1 Subject JD - Phase 1 results

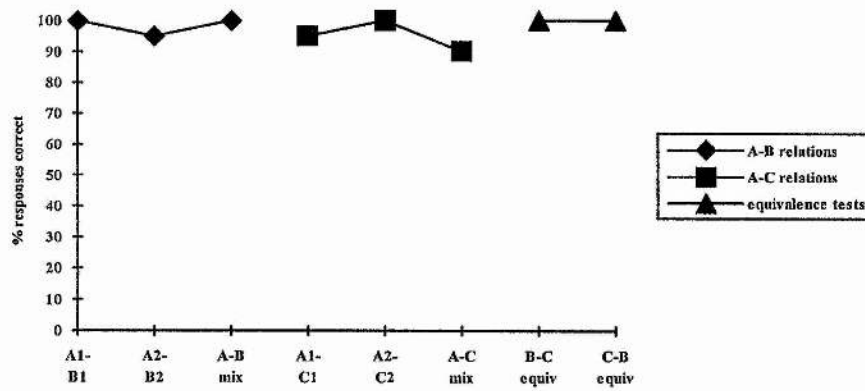


Fig 6:2 Subject TD - Phase 1 results

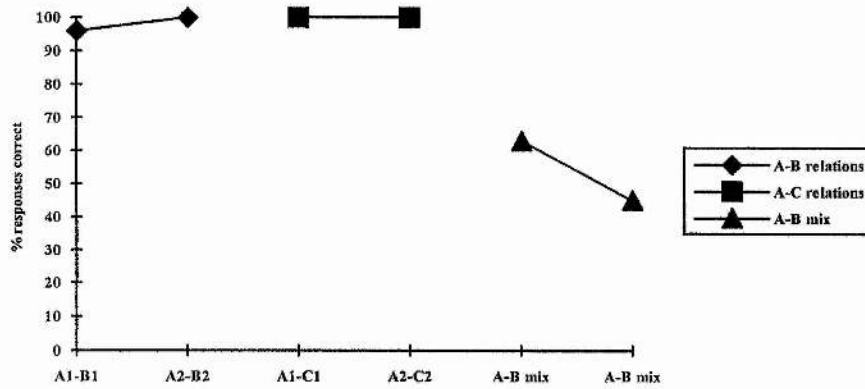


Fig 6:3 Subject TD - Phase 2 results

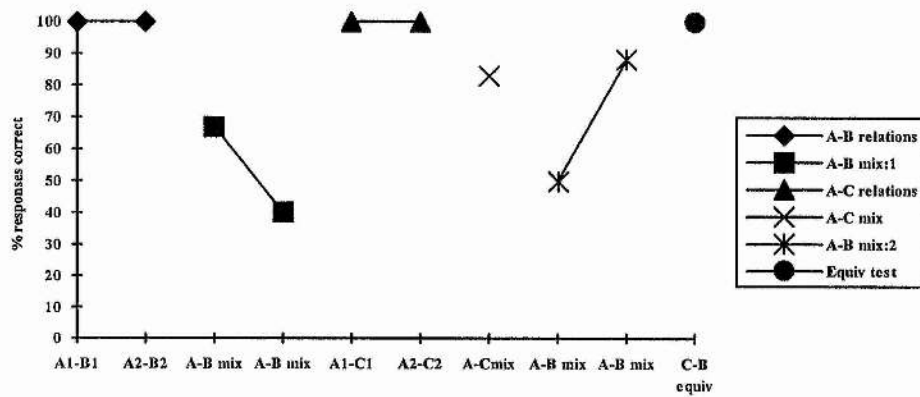


Fig. 6:4 Subject NL - Phase 1 results

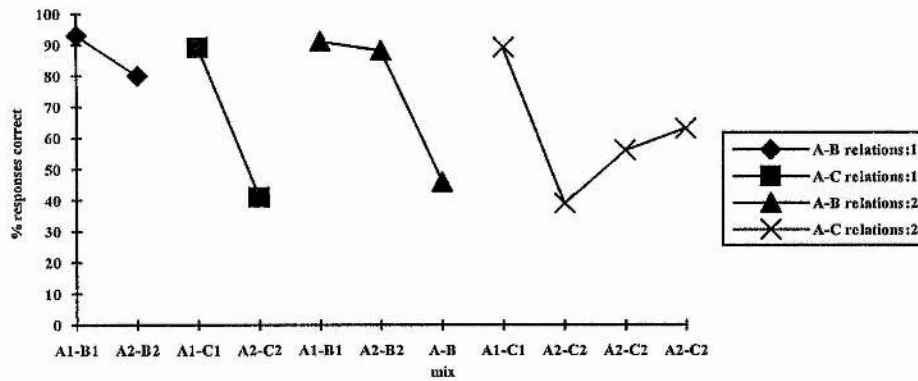


Fig 6:5 Subject NL - Phase 1 results, new equivalence class

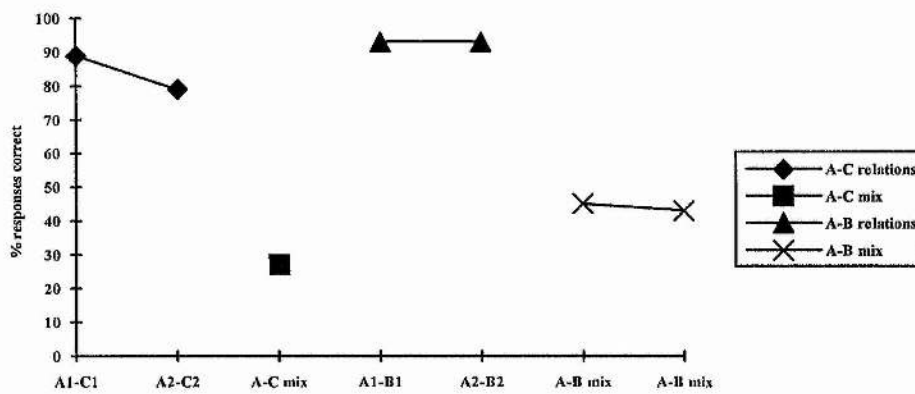


Fig 6:6 Subject NI - Tests for components of equivalence

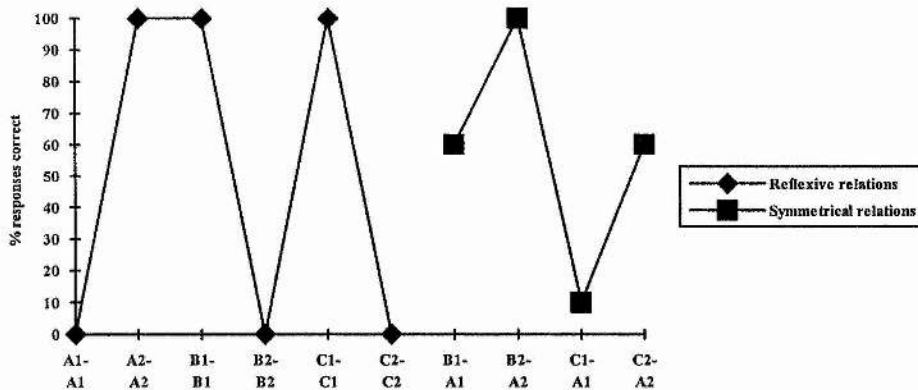


Fig 6:7 Subject GT - Phase 1, trial and error training

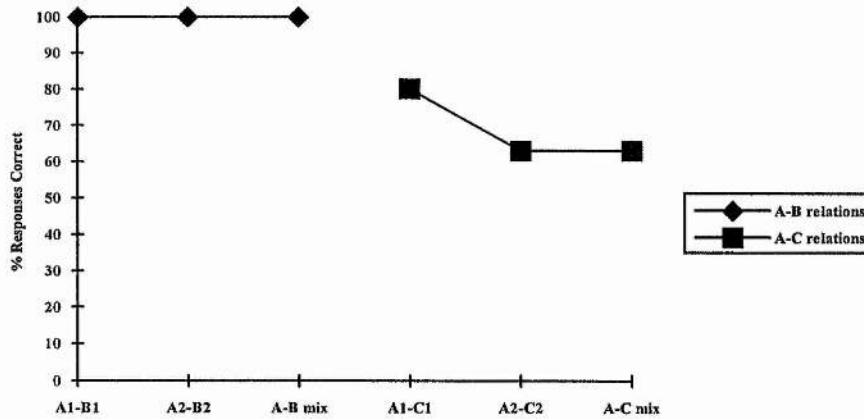


Fig 6:8 Subject GT - Phase 1, errorless learning procedure

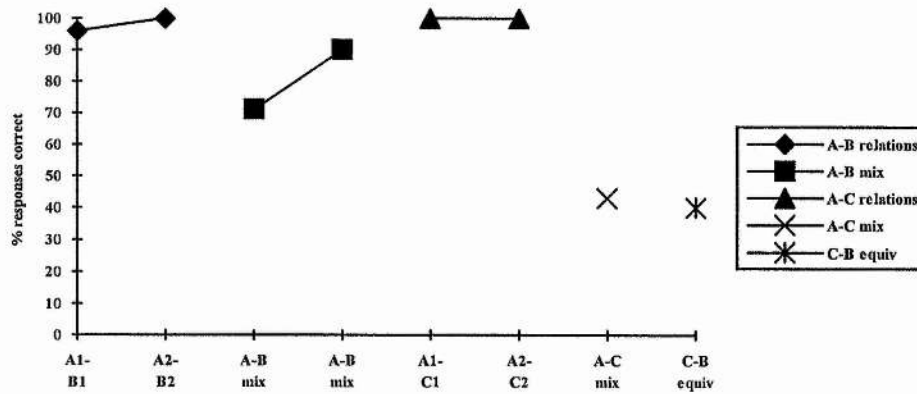
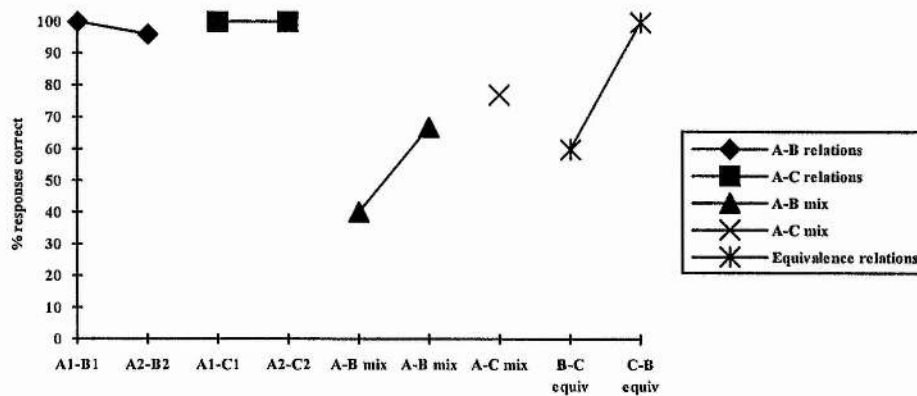


Fig 6:9 Subject GT - Phase 2 results



CHAPTER 7

Experiment 2:

Investigation of Transitive Inference Performance

Stimulus equivalence provides one example of the emergence of untaught relations. Another paradigm which also demonstrates the emergence of new relations on the basis of previously learned relations is transitive inference. In transitive inference, if two stimuli, A and B, are related to each other by relation r , ArB , and stimuli B and C are also related by relation r , BrC , then it is possible to infer that A and C must also be related by relation r , ArC . It seems likely that proficiency on relations such as this underlies the ability to make judgements about the relative size and value of a range of objects and events.

For quite some time, one of the important questions in research on transitive inference was determining at what age normally developing young children are able to display transitive inference skills. Research by Piaget and his co-workers had suggested that children were unable to form transitive inferences until the age of about seven, when they passed the stage of logical pre-operations. Other researchers however, queried this finding on the basis that if this was true, younger children should not be able to understand the most elementary principles of measurement, which would have important educational implications. As Halford (1984) suggested, limitations on children's comprehension of transitivity have implications for virtually all their quantitative thinking.

One possibility was that children were in fact able to make inferences at ages younger than seven, but demonstration of these abilities was confounded by the means with which they were tested. Piaget and his co-workers had tended to use tasks such as three-term transitive syllogisms to examine transitive inference responding, e.g.:

1. Edith is fairer than Suzanne
2. Edith is darker than Lili
3. Who is the darkest, Edith, Suzanne, or Lili ?.

This task requires quite complicated sentence encoding in order to co-ordinate the pieces of information and make an inference. Piaget also used transitive inference tests in which the task itself was not linguistically based. However, on these tasks the subject was required to give a correct verbal explanation of their solution to the problem before they were classed as having made a transitive inference. Thus it

was possible that children younger than seven could in fact make transitive inferences but this was not being demonstrated because of the manner in which these abilities were tested.

Bryant and Trabasso (1971) devised a new procedure for testing transitive inference ability, which did not require complicated linguistic abilities. This involved training a series of four overlapping pair discriminations and then testing performance on untrained non-adjacent pairs of stimuli. This involved presenting a series of coloured wooden rods of different lengths, in pairs, and asking the subject which rod was longer. The rods were presented embedded in a wooden block so that the subjects could not see which rod was longer, but instead had to remember the feedback they were given and integrate the premise information in order to make a transitive inference.

Five rods were used to establish the relations $A > B$, $B > C$, $C > D$, $D > E$. The test for transitive inference was the derivation of $B > D$. Both stimuli B and D were involved in comparisons where they were sometimes longer than the other stimulus and sometimes shorter. This prevented the subjects from performing correctly on the transitive inference test by "parroting" absolute labels. It was also possible to derive $B > D$ by co-ordinating the relational information from the stimulus pairs. Using this technique Bryant and Trabasso were able to demonstrate transitive inference responding with children as young as four years old (see chapter 4 for fuller discussion).

A further advantage of Bryant and Trabasso's procedure was that it was easily adaptable for use with non-human subjects. Using adaptations of this procedure transitive inference has been demonstrated with squirrel monkeys (McGonigle and Chalmers, 1977), chimpanzees (Gillan, 1981), pigeons (Fersen, Wynne, Delius and Staddon, 1991; Higa and Staddon, 1993) and rats (Davis, 1992). (See chapter 3 for fuller discussion).

These findings of transitive inference with non-human subjects seemed to display a significant difference between the paradigms of stimulus equivalence and transitive inference. Both require the derivation of new relations on the basis of explicitly established relations, but it would appear that non-humans will demonstrate transitive inference fairly readily, while it has proved very difficult to conclusively demonstrate stimulus equivalence with non-humans. This then seemed to suggest that behaviourally, transitive inference may be a less complex process than

stimulus equivalence. The fact that non-humans display transitive inference responding suggests that it is not linguistically dependent, as it has been suggested stimulus equivalence may be.

In this case it is possible that subjects who fail to display transitive inference would also fail to demonstrate stimulus equivalence, while subjects who display stimulus equivalence would also be expected to pass tests of transitive inference ability. It would also be possible for a subject to fail a stimulus equivalence test but still pass a transitive inference test. Transitive inference then seemed to make a good comparison test when considering subjects' performances on stimulus equivalence tests. If the two paradigms do appear to be tapping into similar behavioural processes, transitive inference may give some insight into the processes underlying derived equivalence relations.

Two of the subjects who took part in this experiment had previously been given tests of stimulus equivalence. One had clearly demonstrated the formation of equivalence classes, while the other subject had failed to reliably demonstrate any of the properties of equivalence relations. The aim was to see how these subjects' performances on transitive inference tasks compared to their performance on stimulus equivalence tasks, and if this performance matched the predictions about the relation between equivalence and inference performances. It would be predicted that JD who had clearly demonstrated the emergence of equivalence relations would also demonstrate transitive inferences. Subject NL might be able to demonstrate transitive inference although she had failed to demonstrate reliable performance on the components of equivalence (see Chapter 6). According to Bryant and Trabasso (1971), performance by all subjects should be directly related to how securely they learned the baseline pair comparisons.

There were other reasons for carrying out this experiment with adults with learning disabilities as subjects. The majority of research into transitive inference had used normally developing young children, or non-humans as subjects. It is assumed that normally able adults make basic transitive inferences (such as those described) so easily that it is difficult to examine the processes operating from their performances. It is expected that these abilities are less well developed with young children and non-human subjects, and that examining the performances of those on the borderline between success and failure on these tasks will give more insight into the processes operating. In many cases this assumption seems valid. Bryant and Trabasso had hypothesised that a major problem for young children in deriving

transitive inferences was that they could not remember the relational information they needed to combine. Bryant and Trabasso believed that performance on transitive inference tasks was directly related to retention of the baseline comparisons. The results from Bryant and Trabasso (1971) supported this hypothesis. The performance of adults with learning disabilities seems relevant to this proposition.

There is an extensive research literature on the performance of adults with learning disabilities on stimulus equivalence training tasks. On numerous occasions, subjects have demonstrated accurate performance on the conditional discrimination baseline relations, yet still failed to demonstrate emergent equivalence relations. the demonstration of emergent relations such as transitive inference and stimulus equivalence may require more than simply secure acquisition of the baseline relations. In this case it would be valuable to know more about transitive inference performances by subjects who may have specific difficulties in demonstrating these emergent relations.

Method

Subjects

The subjects were three adults with learning disabilities, two female subjects, JD and NL, and one male subject, MM. The subjects were aged between 49 and 70 years old (see subject data, chapter 6). Subjects JD and NL had previously taken part in experiment 1, which tested performance on a stimulus equivalence task. JD had clearly demonstrated the formation of equivalence classes in phase 1 of the experiment. NL did not display equivalence responding, and when tested, did not show reliable performance on any of the properties of equivalence relations.

Stimuli

The stimuli used were similar to those used by Bryant and Trabasso (1971). Five wooden rods of differing lengths and colours were used to teach the overlapping pair discriminations (see Table 7:1). These rods were presented to the subjects embedded in a round wooden block 18 cm's high. Additional pieces of wooden rod were inserted in the holes in this block so that each rod protruded from the top by 3 cm's.

Procedure

The training and test procedure was based on the procedure used by Bryant and Trabasso (1971). The five wooden rods were used to teach a series of four overlapping pair discriminations, $A > B$, $B > C$, $C > D$, $D > E$. The rods were presented to the subjects, embedded in the block, in pairs. The subject was told that one rod was longer than the other, although they couldn't see the difference as the rods were hidden in the block. The subject was told he or she had to try and guess which rod was longer, and to point to the one they thought was longer.

Initially the pairs were taught in order, i.e. $A > B$ was trained first to a pre-set criterion, then $B > C$ was taught to the same criterion, then $C > D$ etc. On this first phase, the criterion was set at five consecutive correct responses on each pair. When the subjects had reached criterion on all four pairs of this task in a single session, they moved on to the next phase of the training procedure. In this second phase, the four pair discriminations were presented to the subject mixed in a random order. Again, the subjects were asked to point to the rod which they thought was longest. Criterion on this task was set at 70% correct responding on each pair, in a single session where each pair was presented at least 10 times. On some occasions, when the subject was first trained on the random presentation of the pairs, the subject's performance deteriorated substantially. On these occasions, the pair discriminations were once again trained in order to try and re-establish stable performance on each pair. The mixed sequence was then reintroduced.

In these two training conditions, the subjects received feedback on their selections, either being told "Yes, that's right", or, "No, the other one is longer". At the same time, the rods were removed from the block and shown to the subject to provide visual feedback as well.

Once the subject had reached criterion on the second phase of the training procedure, the test trials were given. In these trials, the test pair $B > D$ was presented mixed with the other training pairs in the random baseline sequence. Thus, responding on the test pair required the same manner of responding as on all the training pairs. The five training and test pairs were each presented to the subject 10 times in the test session. The subjects received no feedback on the test trials. Instead, they were told that this time they would not be told if they were right or wrong, but they were just to try and point to the rod they thought was longer as well as they could.

Results

Subject JD

JD reached criterion on the first session where the discriminations were presented in order (see Fig. 7:1). When the discriminations were presented mixed in a random sequence her performance was a little unstable so, after three sessions of random presentation, the discriminations were presented in order once more to try and stabilise performance (see Fig 7:2). After two more mixed sessions were performance was slightly improved, there was an unavoidable break in testing for one week. Following this JD reached criterion in one session both when the discriminations were presented in order, and in the subsequent session when the mixed sequence was presented (see Fig 7:3). JD was then given the mixed sequence including test probes. JD maintained performance on the baseline relations and demonstrated perfect transitive inference performance on the B>D test pair.

Subject MM

MM reached criterion on the first session where the discriminations were presented in order (see Fig. 7:4). In the next two sessions when the mixed sequence was presented MM's performance was consistently better on the end pairs of A>B and D>E than on the middle pairs of B>C and C>D. The discriminations were then presented in order once more to try and improve performance on the middle pairs before the random sequence was presented once more.

This pattern was repeated several times (see Fig. 7:5); performance was good on all pairs while the discriminations were presented in order, but on the mixed sequences performance was consistently better on the end pairs than the middle pairs. This was probably exacerbated by two unavoidable one week breaks, between sessions 5 and 6 and between sessions 7 and 8.

In an attempt to remedy this, on session 13 the two middle pairs were trained in order without the end pairs (see Fig. 7:6). This was intended to selectively improve performance on these pairs before re-introducing the random sequence. This appeared to have some effect as performance on the mixed sequence in the following two sessions was much improved. MM actually reached criterion on the mixed sequence in the second of these two sessions (session 15). However, as MM's performance had been so unstable previously, one further mixed sequence was given to try and ensure performance was stable before testing for equivalence.

On this session (session 16) MM again showed a marked deterioration of performance on $C > D$.

To try and correct this, the procedure of training the $B > C$ and $C > D$ relations separately and then reintroducing the mixed sequences was repeated twice more (sessions 17 to 20) (see Fig. 7:6 and Fig. 7:7). This resulted in stable performance on all pairs in a random sequence (sessions 20 and 21) and MM was given the test probes (see Fig. 7:7). During the test session, MM demonstrated substantial disruption of the $B > C$ pair, with responding on this item seeming to reverse. MM also failed to display transitive inference responding. On the test pair $B > D$ MM showed a preference for D over B, seeming to match the reversal on $B > C$.

Subject NL

It proved very difficult to achieve stable performance with subject NL (see Fig. 7:8). Even when the discriminations were presented in order NL did not reach criterion on all four pairs in one session until the fifth training session. This was not helped by a two week break in testing between sessions 3 and 4. When the discriminations were presented in a mixed sequence (session 6), NL showed disruption of performance on the last pair in the sequence $D > E$. On the next session, this disruption seemed to be affecting the adjacent $C > D$ pair as well. To try and correct this the discriminations were presented in order again (session 8, see Fig. 7:9), and the random sequence was then reintroduced.

Performance remained poor on the $D > E$ pair, with the discrimination almost reversing, even when performance on the other pairs was perfect (session 11). NL was then given training on $D > E$ alone (session 12), and the mixed sequence was reintroduced. However, this appeared to move the disruption to the $C > D$ pair. Over the next 12 sessions, NL was given training on the discriminations in a cycle of training, first in order, and then in a random sequence (see Fig. 7:10 and Fig 7:11). Despite this NL never maintained sufficiently stable performance to be tested for transitive inference. Typically performance was below criterion on the $D > E$ pair, even when the discriminations were trained in order.

Discussion

Of the three subjects tested in this experiment, one subject (JD) clearly demonstrated transitive inference responding, while the other two subjects (MM and NL) did not. Subject MM failed to display transitive inference when given the

test probes while NL never displayed sufficiently stable performance on the baseline relations to be tested for transitive inference responding.

MM's failure to demonstrate transitive inference is predictable given the deterioration of performance on one of the baseline relations during testing. Bryant and Trabasso (1971) had suggested that the failure of young children to display transitive inference may have been due to their failure to remember the comparisons they needed to combine. This is one reason why performance on the baseline relations is tested at the same time as the inferential $B > D$ test pair. Bryant and Trabasso showed that the probability of making a correct inferential judgement on the test pair is the probability of jointly recalling the information for each of the training pairs. Therefore, disruption of performance on any of the training pairs makes it less likely that correct inferential performance will be demonstrated. In MM's case performance not only deteriorated but in fact seemed to reverse, making it even less likely that $B > D$ performance would be demonstrated. As $B > C$ had reversed, MM was no longer combining the information $B > C$, $C > D$, but rather $C > B$, $C > D$. Thus stimuli B and D had effectively equalised in value, making a true transitive inference impossible. MM may have consistently selected D rather than B on test trials, as he now had a history of selecting D in the pair $D > E$, while B was no longer preferred over any of the other stimuli - $A > B$, " $C > B$ ".

This deterioration in MM's performance may be due to the sudden withdrawal of feedback for responding. Until the final test trial MM had received feedback on all training trials, but on the test block no feedback was given at all. Possibly, if reinforcement had been gradually reduced over several training sessions MM would not have shown this disruption of baseline performance when given the test pair.

Subject JD not only demonstrated appropriate transitive inference performance, but actually showed superior performance on the $B > D$ test pair than on the four baseline training relations. This sort of performance has been noted before in the inference literature and is known as the Symbolic Distance Effect. This implies that it is easier to judge the size or value of stimuli the further apart they lie on some form of relative scale. "Thus the time taken to judge the relative size of, e.g. "cat" vs. "whale" is faster than that required to determine the relative size of "cat" vs. "fox"" (McGonigle and Chalmers, 1984, p.525). It has been suggested that the training procedure used in this experiment establishes the linear series $A > B > C > D > E$, and that subjects make judgements about the relative value of

stimuli in this series by "reading off" their relative positions on this line. If this is true, then $B > D$ is separated by one stimulus (or node) while the baseline discriminations are all adjacent to each other. Thus superior performance on $B > D$ would be predicted by virtue of the greater symbolic distance between the stimuli. Theoretically, performance on the test pair $A > D$ would be expected to be even better than on $B > D$ as A and D are separated by two stimuli. However, as selection of A is always correct compared to any of the other stimuli in the series, any results on this task would have been confounded by absolute responding.

The results from this experiment generally confirm the predictions about relative performances on stimulus equivalence and transitive inference tasks. NL failed to demonstrate transitive inference in this experiment and had previously failed to demonstrate stimulus equivalence. Conversely, JD had demonstrated stimulus equivalence in experiment 1 and, as expected, demonstrated appropriate inference performance in this experiment. However, as only three subjects took part in this experiment it is hard to conclusively confirm or disprove the predictions about equivalence and inference performance. The performance by JD is more instructive in this case than that of NL. JD clearly demonstrated both equivalence and inference performance in these experiments, but NL was never explicitly tested for either of these relations. Instead NL is assumed not to have demonstrated either of these properties as she did not reach criterion on the baseline training relations in either experiment. However, as Bryant and Trabasso (1971) demonstrated, a subject cannot be expected to demonstrate relations of transitive inference if they cannot remember the necessary relational premises. It would still be interesting, though, to examine NL's performance if it was possible to find some way to reliably teach the baseline relations for either experiment. NL clearly has difficulty achieving stable performance on the baseline trials. If some way could be found to teach the necessary relational information, would NL then be able to use this information to make a transitive inference, or to derive equivalence classes.

A similar argument could be applied to the performance of MM on this task. It is not clear if MM is able to make transitive inferences or not. MM did fail the transitive inference test, but this is predictable due to the deterioration in his baseline performance. If gradual phasing out of feedback over several sessions had served to maintain this baseline performance, it is possible that MM would have been able to derive $B > D$. As it appears that MM displays transitive inferences less reliably than JD, it would be interesting to see if MM could display stimulus

equivalence if given conditional discrimination training, or if the two paradigms are in fact at different points on a continuum of relational properties.

Sidman (1994) has suggested that the derivation of stimulus equivalence relations is a fundamental stimulus function which results from reinforcement contingencies. A similar argument can be made for the relations demonstrated here. Several transitive inference studies (Fersen, Wynne, Delius and Staddon, 1991; Gillan, 1981; McGonigle and Chalmers, 1977) have characterised the reinforcement given on the baseline training pairs as giving one stimulus a value of 1, and the other stimulus a value of 0. For example, on $B > C$, B would acquire a value of 1 and C a value of 0. It is easy to see how the reinforcement contingencies in training could establish the stimulus series $A > B > C > D > E$ and thus the relative values of the test pair $B > D$. Thus, the demonstration of transitive inference could be explained by reinforcement contingencies in the same way that Sidman suggests stimulus equivalence can be explained. This might also explain why the sudden withdrawal of reinforcement had such an effect on performance by MM.

If this is true, it would suggest even more strongly that the demonstration of both stimulus equivalence and transitive inference are part of a more general ability to derive emergent relations on the basis of previously learned information. This ability may be part of a fundamental stimulus function; a product of survival contingencies as Sidman suggests. After all, it has been suggested that "the ability to rank objects on a hedonic scale and make judgements about the desirability of items that have never been encountered together must often have adaptive value" (Fersen, Wynne, Delius and Staddon, 1991, p. 334).

Table 7:1 Training and Test Stimuli

Stimulus	A	B	C	D	E
Colour	Blue	Red	White	Yellow	Green
Length (cms)	18	15	13	10	8

Fig. 7:1 Subject JD - Training Sessions 1 to 4

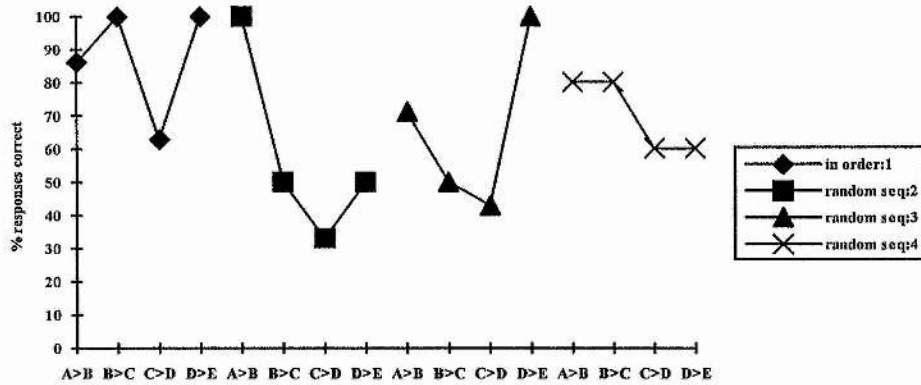


Fig. 7:2 Subject JD - Training Sessions 5 to 7

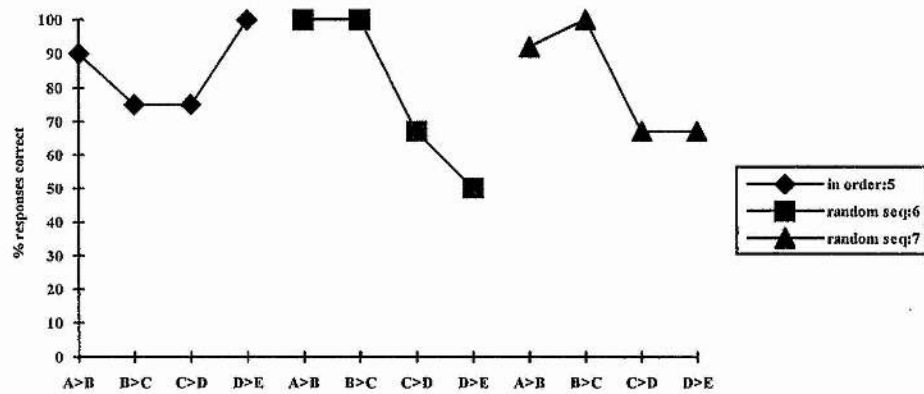


Fig. 7:3 Subject JD - Training and Test Sessions 8 to 10

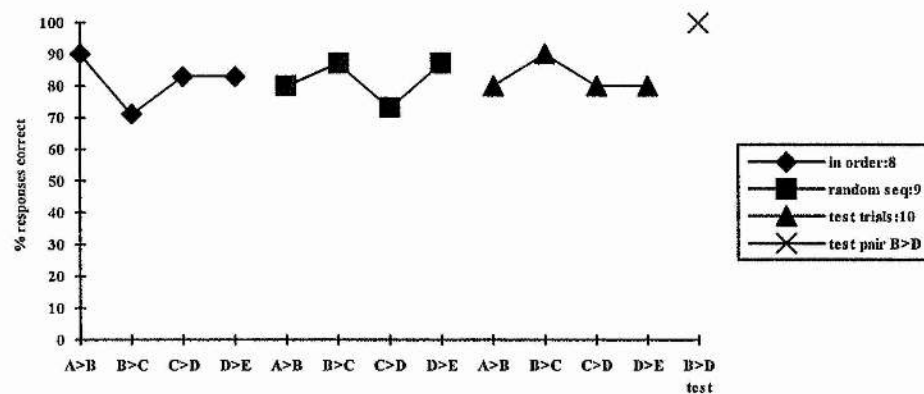


Fig. 7:6 Subject MM - Training Sessions 13 to 18

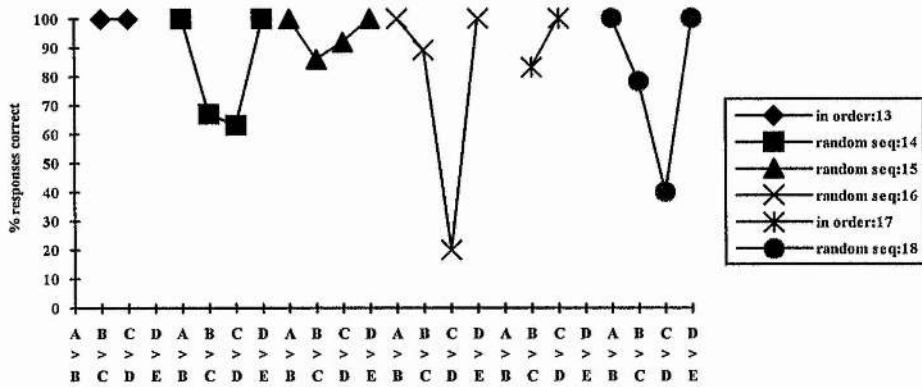


Fig. 7:7 Subject MM - Training and Test Sessions 19 to 22

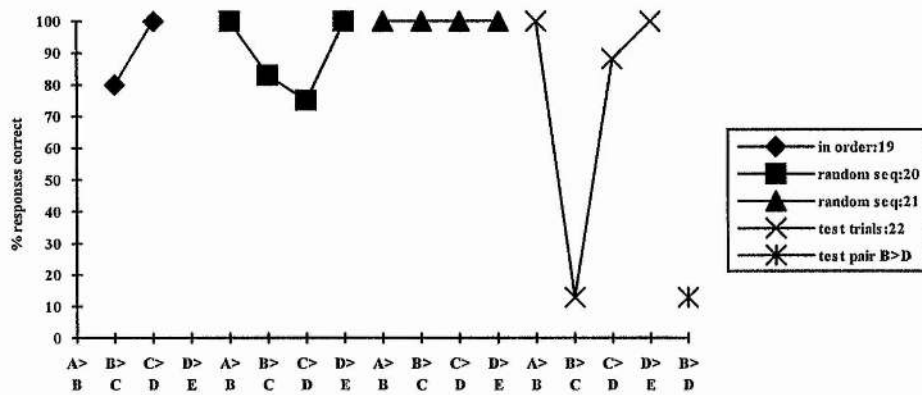


Fig. 7:8 Subject NL - Training Sessions 1 to 7

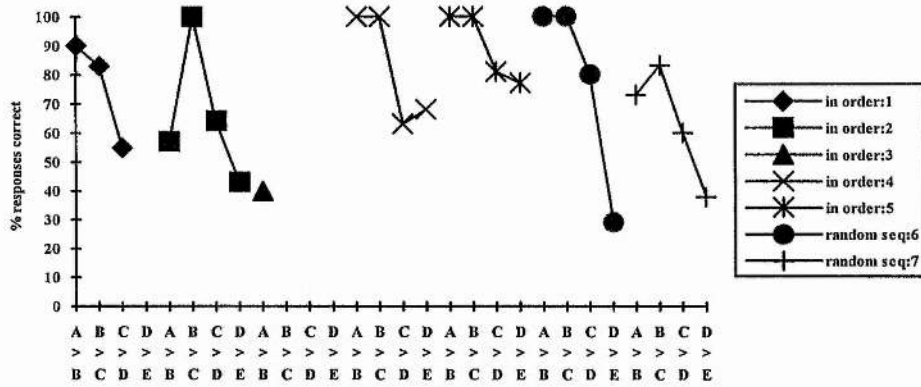


Fig. 7:9 Subject NL - Training Sessions 8 to 13

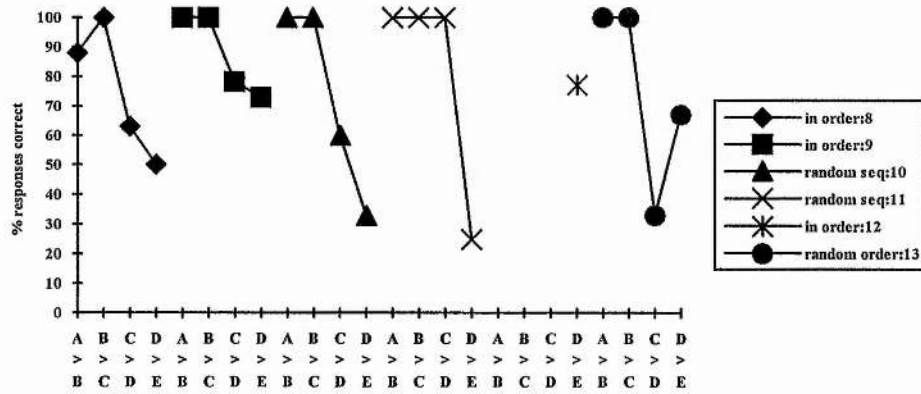


Fig. 7:10 Subject NL - Training Sessions 14 to 19

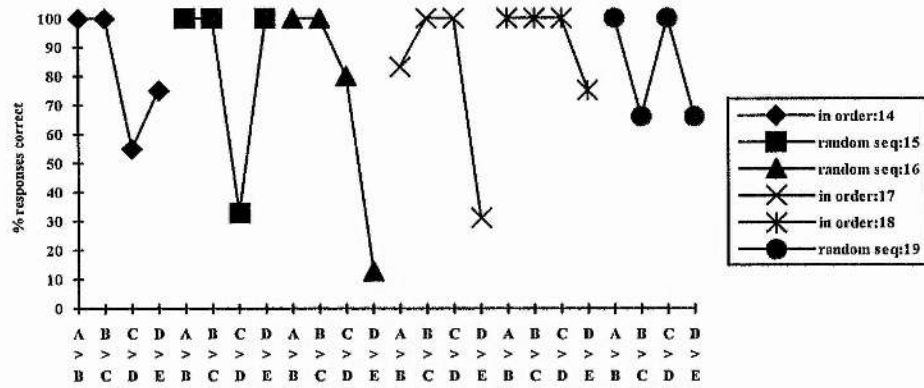
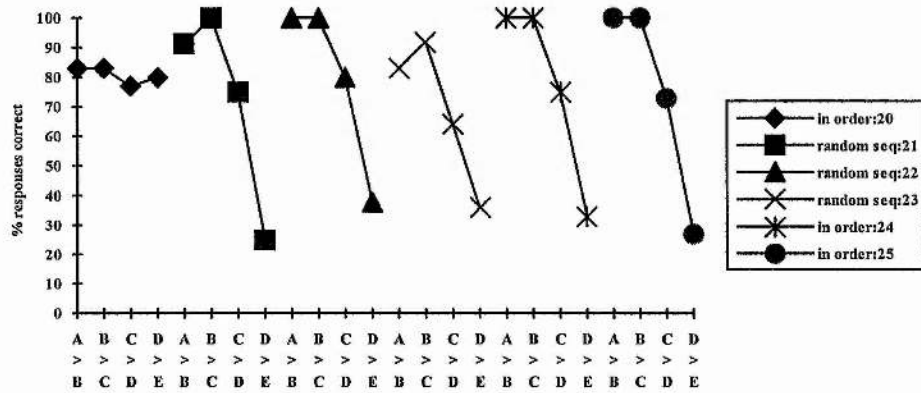


Fig. 7:11 Subject NL - Training Sessions 20 to 25



CHAPTER 8

Experiment 3:

Disruption of Transitive Inference Performance

One reason for studying performance on transitive inference tasks is that, like equivalence, it involves the emergence of untaught relations from previously established relations.

"Transitive Inference denotes the ability to infer relationships between items that have not been trained together" (Fersen, Wynne, Delius and Staddon, 1991)

Transitive Inference had been studied originally using tasks such as three term syllogisms (cf. Piaget, 1928). Bryant and Trabasso (1971) had argued that these tasks were too language dependent and had developed an alternative procedure. They used five wooden rods of differing lengths and colours to train four overlapping pair discriminations, $A > B$, $B > C$, $C > D$, $D > E$, and tested for accurate performance on non-adjacent pairs, in particular $B > D$, the most stringent test for transitive inference. Using this procedure Bryant and Trabasso demonstrated transitive inference performance in children as young as four.

One advantage of this procedure developed by Bryant and Trabasso was that it was fairly easily adapted for use with non-human subjects. Using a procedure based on Bryant and Trabasso's, McGonigle and Chalmers (1977) were able to demonstrate transitive inference with squirrel monkeys. This finding was subsequently replicated with a number of other non-human subjects (Davis, 1992, Fersen et al, 1991, Gillan, 1981, Higa and Staddon, 1993; see Chapter 3 for fuller discussion).

Thus transitive inference requires the demonstration of emergent untaught relations, and has been demonstrated in a number of non-human species. Stimulus equivalence also requires the demonstration of emergent untaught relations, but has proved difficult to demonstrate with non-humans (Dugdale and Lowe, 1990, Hayes, 1989, McIntire, Cleary and Thompson, 1987, Schusterman and Kastak, 1993, Vaughan, 1988; see Chapter 3 for fuller discussion). By examining the differences between the two paradigms in this respect it might be possible to gain some insight into the processes underlying emergent relations.

One way of examining the necessary conditions for transitive inference might be to establish transitive inference responding, and then to disrupt one aspect of the

baseline relations and see what effect this has on the other relations. The rationale for this comes from a study carried out by Gillan (1981) examining transitive inference with chimpanzees. Using a procedure based on Bryant and Trabasso's (1971) procedure, Gillan trained four overlapping pair discriminations, $A < B$, $B < C$, $C < D$, $D < E$, and tested for the emergence of the transitive inference relation $B < D$. Having established transitive inference, Gillan was able to extend the series by the addition of another training pair $E < F$, demonstrating transitive inference responding on two new test pairs, $B < E$ and $C < E$. This transitive inference responding was destroyed by training a further relation, $F < A$, which closed the series by removing the end points. Transitive inference responding was restored by subsequently training $A < F$, which restored the end points of the series, making it consistent with the original training. These findings were subsequently replicated by Fersen et al (1991) with pigeons, and Davis (1992) with rats. From this work Gillan proposed that there were two necessary conditions for correct choice on transitive inference tasks:

1. Accurate performance on the adjacent training pairs.
2. The stimuli that form the series must be ordered on some uni-dimensional scale.

A study by Siemann and Delius (1994) carried out a similar manipulation. With normally able human adults they used 10 overlapping pair comparisons to establish stimulus sequences. For one group (G0) this sequence complied with a linear structure forming the series $A > B > C > D > E > F$. For another group (G1) one of the training pairs was inconsistent with this linear sequence, establishing an "odd" series, while for a third group (G3), three of the pairs were inconsistent with the linear sequence, instead forming a circular series overall. Siemann and Delius found that all the subjects were able to respond on the pairs making up the series with much the same accuracy, but the response latencies for each group were significantly different, increasing from G0 to G3. When a similar manipulation was carried out with pigeons, they showed increasing disruption of performance from G0 to G3 (see Chapter 4).

These results seemed to replicate previous findings of disruption of transitive inference performances with non-human subjects. However, one important difference was that the human subjects were able to cope with even the most difficult task, G3, albeit at the expense of long reaction times. Non human subjects tend to respond at around chance level, or cease responding altogether, when the linear series is disrupted.

This study by Siemann and Delius was the first to report a manipulation of this kind with human subjects. It seemed likely that the subjects they used, normally able adult humans, would be best able to maintain responding on a disrupted transitive inference series. However, it is not clear how well subjects such as normally developing young children, or adults with learning disabilities would cope with such a manipulation. While these subjects can demonstrate transitive inference responding, their performance may not be as robust as that of the subjects in Siemann and Delius (1994). Examination of how these subjects cope with such a manipulation might provide some insight into the conditions necessary for accurate transitive inference performance.

There is a considerable literature on performance by normally developing young children on standard transitive inference tasks. It is harder to predict how they would cope with a task which disrupts what appear to be the necessary preconditions for transitive inference. Bryant and Trabasso (1971) successfully demonstrated that children as young as four could make transitive inferences. However, Kuczaj and Donaldson (1982) demonstrated that there may be a developmental pattern to the way in which children learn to make these inferences. Thus, while young children may be able to make transitive inferences their inferences may not always be appropriate. The manipulation described in Gillan (1981) of disrupted the series by removing the end points. The normally able adults in the study by Siemann and Delius (1994) were able to cope with this manipulation, but the significantly increased response latencies suggests that they found it difficult. Young children who have not yet achieved a fully "adult" ability to make transitive inferences may not have the skills necessary to maintain responding. Testing young children on this task may give some insight into the developmental pattern of transitive inference ability and help clarify some of the conditions necessary for successful transitive inference performance.

Much less is known about performance on transitive inference tasks by adults with learning disabilities. Experiment 2 (Chapter 7) described one attempt to demonstrate transitive inference with this group. One subject clearly demonstrated transitive inference performance, another subject failed a transitive inference test while also showing loss of performance on the training pairs, and a third subject did not achieve sufficiently stable performance to be tested.

The ability to make transitive inferences is likely to be important when working with concepts such as quantity and relative value and for many of the skills

involved in basic mathematics. However, these are things which many adults with learning disabilities have specific problems with. The normally able adults in the study by Siemann and Delius (1994) dealt with the manipulation of the training pairs. It is possible that adults with learning disabilities may not show such robust transitive inference performance. Just as Gillan's (1981) manipulation gave some idea of the basic pre-requisites for transitive inference performance, so a similar manipulation with learning disabilities may give some idea of the specific problems encountered as transitive inference tasks become increasingly complex. After all, it is most unlikely that all relational information forms a neatly ordered series such as that encountered in transitive inference tasks.

In this experiment, transitive inference performance was established with three different groups of human subjects, normally able adults, an adult with learning disabilities, and normally developing young children. Once transitive inference responding was established the training series was disrupted by training one inconsistent pair, and transitive inference responding was then reassessed.

Method

Subjects

Three different groups of subjects were tested:

1. Normally able adults.

These subjects were 10 undergraduate students recruited by means of a sign up sheet in the psychology department. Participation was limited to first year psychology students, or non-psychology students, to ensure that the subjects were not familiar with the research area. The subjects were paid £3 an hour for their participation.

2. Subject with learning disabilities

The subject was MM, a male aged 50. MM had previously taken part in an experiment on transitive inference (see chapter 7, and subject data chapter 6) but had failed to demonstrate transitive inference responding. MM was paid for his participation, receiving 10p every 5 trials.

3. Normally developing young children

These subjects were eight children aged between 5:1 and 5:8 (mean age 5:4; years:months). Four of these subjects did not have English as their first language,

but all were competent in English and all the children were normally taught in English.

Stimuli

The stimuli used were the same for all subjects. These were five wooden rods of differing lengths and colours (see Table 8:1). These were presented to the subjects embedded in a wooden block so that only the top 2 cm's protruded. A screen was used to conceal the block and wooden rods while each trial was being prepared. Additional stimuli used only with the young children were a tray with 40 neutral coloured wooden beads and a small glass jar.

Design

The basic design of the experiment was the same for all subjects, and consisted of two phases.

Phase 1

The aim of phase 1 was to establish stable transitive inference responding. The five wooden rods were used to train a series of four overlapping pair discriminations, $A > B$, $B > C$, $C > D$, $D > E$.

Initially these discriminations were trained in order, first $A > B$ to criterion, then $B > C$ etc. When the subjects reached criterion on this task, the four discriminations were presented mixed in a random order. When the subjects reached criterion on this task they were given the test probes. In these trials the transitive inference test pair, $B > D$, was presented mixed with the baseline training pairs in the random baseline sequence. This tested for retention of the baseline relations and selection of B rather than D, the test for inference. Provided the subjects demonstrated appropriate transitive inference performance they progressed to phase 2 of the study.

Phase 2

The aim of phase 2 was to train one further pair discrimination, $E > A$, which effectively made the baseline circular, destroying the linear order of the series. The subjects were then tested on three non-adjacent stimulus pairs.

As in phase 1, initially the discriminations were trained in order, i.e. $A > B$, then $B > C$, $C > D$, $D < E$, and then the new pair $E > A$. When the subjects reached criterion on this task, the five discriminations were presented mixed in a random sequence.

It was possible that the addition of the $E > A$ discrimination might cause disruption to previously stable performance on the other discriminations. For this reason a specific criterion was not set for this mixed baseline. Instead, the amount of training each subject received on this task was matched to the amount of training they had received on the equivalent task in phase 1. The subject was then given the test probes. This consisted of the five baseline discriminations and three non-adjacent test pairs. These were, the transitive inference test pair from phase 1, $B > D$, and two other pairs $B > E$ and $A > D$. The $B > D$ test pair was retained to see if training of the inconsistent relation $E > A$ would affect responding on this pair. The two new pairs were included to assess how the subjects would deal with non-adjacent pairs now that the stimuli no longer formed a linear series.

In both phases, feedback was given on the subjects' selections when the discriminations were trained in order, and on the mixed baseline training trials. On the test probes, no feedback was given on the baseline pairs or the test pairs.

Procedure

The procedure for training and testing varied slightly for each group of subjects:

1. Normally able adults

These subjects received the basic training procedure outlined above. Before training the subjects were told they would be shown a block with two rods inserted in it. They were told that the rods were different lengths, but that the differences were concealed by the block, and that their job was to try and decide which rod was longest. They were also told that sometimes they would be told if their selection was correct or incorrect, and that sometimes they would not, but that they should try and respond as best they could on all trials.

Phase 1 - Initially the rods were presented in order, to a criterion of five consecutive correct responses on each pair. They were then given the random baseline sequence. Criterion on this task was set at 80% correct responding on each pair, or if it took some time for performance to stabilise, the last five responses on each pair had to be correct. As long as one of these criteria were met, the subjects received the test probes. In the test probes, the subjects were tested on the four baseline discriminations and the $B > D$ test pair. No feedback was given on this task.

Phase 2 - As in phase 1, the five baseline discriminations were initially trained in order, to a criterion of five consecutive correct responses. The subjects were then

trained on the random mixed baseline. No mastery criterion was set for this task, but instead each pair was presented 10 times, as in phase 1. The subjects then received the test probes. These consisted of the four baseline discriminations $A > B$, $B > C$, $C > D$, $D > E$, $E > A$, and three test pairs, $B > D$, $B > E$, and $A > D$. No feedback was given during testing.

Both phases 1 and 2 were carried out in one session, which lasted just over one hour. The subjects were paid for their participation.

2. Adult with learning disabilities.

The basic format of MM's training and testing was the same as for the normally able subjects, but the sessions were shorter and training and testing for both phases took place over a period of 10 days. Before training began, MM was given the same instructions as those given to the normally able adults.

Phase 1 - Initially MM was taught the four pair comparisons, in order, to a criterion of five consecutive correct responses on each pair. After this the four comparisons were presented in random order to a criterion of 80% correct responding on each pair in a single session. Two sessions were then spent gradually reducing the feedback given, in preparation for the test trials. MM had been given transitive inference training on a previous occasion, but had failed to demonstrate transitive inference (see Chapter 7). It is possible that this failure may have been due to the sudden absence of feedback when given the test trials. In this failed attempt, performance on the training relations had been at 100%, but in the absence of feedback MM showed disruption on these relations, as well as failure to demonstrate transitive inference. On this occasion, feedback was gradually reduced to occurring on 20%-25% of trials by the test session. MM was then given the test probes. During the test trials feedback was maintained at 25% of training pairs, but no feedback occurred on the test items.

Phase 2 - In phase 2, the new pair $E > A$ was introduced to close the series. The five pairs $A > B$, $B > C$, $C > D$, $D > E$ and $E > A$ were trained, in order to a criterion of five consecutive correct responses. These five pairs were then presented in a random order. The amount of training on this task was matched to the amount of training given in the same task in phase 1. Once again, over several sessions, feedback was gradually reduced to 20%-25% of trials prior to testing. MM was then given the test probes of five baseline pairs and three test pairs. As in phase 1, feedback was maintained at 25% on the training pairs, but no feedback occurred on the test pairs.

MM was paid for his participation in this experiment, receiving 10p every five trials. This occurred during both training and test trials and was not contingent on correct responding on that trial.

3. Normally developing young children.

The basic format of training and testing was the same as for the other two groups of subjects. The sessions were shorter than those for the normally able adult subjects and took place over several weeks.

Some additions were made to the procedure with regard to feedback and reinforcement. In both phases 1 and 2, before all training sessions, a container with 40 wooden beads and a small glass jar were set on the table in front of the subject. Before the first training session the subject was told that they were going to play a game in which they could earn beads, and that every time they made a correct response they would earn one bead. When all the beads in the tray had been earned (put in the glass jar), the subject received a pre-selected back up reinforcer (a small picture). If, during a training session, the subject earned all the beads, the beads were exchanged for the pre-selected picture, the beads were emptied back into the container, and a new picture was selected. The training session then continued as before. Before test and maintenance sessions the subject was shown the glass jar with the beads earned during the previous session. The jar and tray of beads were then removed from the table and the subject was told, "Now we are going to play the game without me telling you if you are right or wrong. You will also not get any beads while we are playing, but I will keep a record of how many beads you earn and you will be given your beads at the end of the game. Do your best". At the end of the test session, the subject was given one bead for every trial completed.

Phase 1 - In phase 1 the subjects were taught the four overlapping pair discriminations, in order, to a criterion of five consecutive correct responses. The discriminations were then presented mixed in a random order in blocks of 20 trials. Criterion for this mixed block of trials was 80% correct responding on each discrimination in a single session. Once the subjects had reached criterion, they were given a session where a random baseline sequence was given, but without reinforcement. This was to ensure that the performance was maintained in the absence of feedback in preparation for testing. Provided the subjects maintained criterion performance they received the test probes. If performance on the mixed training block was particularly poor, with performance on all or most of the discriminations at chance level, the discriminations were once again trained in

order. For some subjects who consistently failed to achieve stable performance on the mixed block of trials, the four discriminations were split into two mixed blocks of two discriminations. Thus, in one session the subject would receive training on a block of 20 trials of mixed $A > B$, $B > C$ pairs. In the next session, that subject received training on a block of 20 trials of mixed $C > D$, $D > E$ training. When performance was above 80% correct on all four discriminations on these "split" blocks, the mixed block was presented once more.

Phase 2 - Training in this phase was very similar to that in phase 1. The four baseline pairs, with the addition of the new $E > A$ pair, were trained in order to a criterion of five consecutive correct responses. These five pairs were then presented in the random baseline sequence, training on this task being matched to training on the equivalent task in phase 1. For testing, the training and test pairs were presented mixed in a random sequence. The eight pairs were presented four times each in a block of 32 trials. These test probes were presented to each subject on three separate occasions, so that each training and test discrimination was tested a total of 12 times. A session of mixed baseline training was given between each test session.

Results

When assessing performance on the test probes in phase 2, responding on the baseline pairs was deemed correct if it was consistent with what had been established during training. There was no "correct" response for the test pairs in this phase, instead the subjects were assessed to see if they responded consistently to one of the stimuli in each pair. The test stimuli have been written in the form $B > D$, $A > D$ and $B > E$. As the series has effectively been made circular, this is not fully appropriate, but it preserves the notation from phase 1. It does not however imply that the subjects were expected to select stimulus A in preference to stimulus D for example.

1. Normally able adults

In phase 1 all subjects reached criterion when the discriminations were trained in order, and when presented in a random sequence (see Table 8:2). When given the test probes, most subjects maintained performance on the baseline discriminations at a very high level of accuracy. Subject 9 reversed the final $D > E$ discrimination, always selecting E rather than D. All subjects showed perfect performance on the $B > D$ test pair and progressed to phase 2.

In phase 2, all subjects reached criterion when the discriminations were trained in order (see Table 8:3). On the random baseline sequence, all subjects maintained highly accurate performance. When given the test probes, most subjects maintained performance on the baseline pairs. Subject 7 showed some disruption, responding at 30% correct on C>D and 70% correct on D>E. Subject 9 again appeared to reverse responding on D>E, scoring only 10% correct on this pair. Responding on the test pairs was generally characterised by consistent responding to one or other of the stimuli in each test pair. Only subject 7 showed some inconsistent responding. On the B>D test pair, this subject selected B 30% of the time and D 70% of the time and on the B>E test pair, selected B 40% of the time and E 60%. Other than that, subjects responded to one or other of the stimuli in each test pair with an accuracy of 80% or above.

On the B>D test pair, 9 subjects consistently selected B rather than D. Subject 7 had shown a slight preference for D over B. On the B>E test pair, 5 subjects consistently selected B over E, 4 subjects consistently selected E over B, with subject 7 showing no clear choice. On test pair A>D, 7 subjects consistently selected A over D and 3 consistently selected D over A (see Fig 8:1)

2. Adult with learning Disabilities

In phase 1 MM reached criterion on the first session when the five discriminations were presented in order (see Table 8:4). He also showed highly accurate performance on the random presentation of the baseline pairs, and maintained this performance as the rate of feedback was reduced. On the probe trials MM maintained performance on the baseline relations, and demonstrated perfect performance on the transitive inference test pair. MM then progressed to phase 2 (see Fig 8:2)

MM reached criterion on all the phase 2 baseline pairs in the first session when they were presented in order (see Table 8:5). When the five pair discriminations were presented in a random sequence, MM showed disruption to previously stable performance. This disruption increased each time the random sequence was presented. During the final test session performance on the baseline pairs was at chance level for four of the five pairs, and MM showed no consistent choice on any of the test pairs (see Fig 8:3)

3. Normally Developing Young Children

The subjects required between 6 and 19 sessions to complete training and testing on phase 1. Three of the eight subjects, subjects 1, 4 and 5, did not achieve stable performance on the mixed baseline sequences, despite considerable amounts of training (see Table 8:6). When tested, subject 4 failed to demonstrate $B > D$ (20% correct), while subjects 1 and 5 did select $B > D$ but showed inconsistent performance on the baseline pairs during testing. The other five subjects generally showed highly accurate performance on the random baseline sequence. On the test probes, all five subjects maintained accurate performance on the baseline pairs and demonstrated accurate transitive inference performance. These subjects progressed to phase 2.

These subjects received between 7 and 11 training sessions and 3 test sessions in phase 2. This training had been matched to that received in phase 1. All subjects reached criterion in one session when the five baseline relations were presented in order. During training on the random baseline sequence, performance generally remained highly accurate. When given the test probes, performance on the baseline pairs again generally remained highly accurate (see Table 8:7). The only exception to this was subject 3, who failed to maintain performance on $E > A$ during testing (58% correct) and who showed some disruption of performance on $C > D$ (67% correct) and $D > E$ (75% correct). Performance on the test probes was generally characterised by consistent responding to one or other of the stimuli in each pair, by each subject. An exception to this was subject 6 on test pair $A > D$ where the subject showed no clear preference, selecting stimulus A 42% of the time and stimulus D 58% of the time. On the other two test pairs, this subject did show a clearer selection, on pair $B > D$ selecting D 67% of the time and on $B > E$ selecting E 67% of the time, but this choice was not as clear as that shown by the other subjects. No overall preference was shown for either of the stimuli in any of the test pairs (see Fig 8:4). On test pair $B > D$, three subjects selected B and two selected D; on $B > E$ two subjects selected B and three selected E; while on $A > D$, two subjects selected A, two selected D, and one subject showed no overall preference.

Reliability

Reliability data were collected with this group of subjects. The experiment consisted of 4088 trials, 3408 training trials and 680 test trials. Reliability checks were made on 441 training trials (12.94%) and 203 test trials (29.85%). The experimenter and observers always agreed.

Discussion

The aim of this experiment was to establish transitive inference and then examine the effects of disrupting the linear order of the baseline series. On the test probes in phase 2, the five baseline discriminations were deemed correct if responding was in accordance with what had been reinforced in training. There was no "correct" answer to the three test probes $B > D$, $B > E$, $A > D$. What was important was whether the subjects would be able to maintain performance on the baseline relations, and would they be able to respond consistently to one or other of the stimuli in each test pair.

The normally able adults fulfilled both these criteria. Even after the introduction of the $E > A$ inconsistent relation, performance on the baseline relations remained highly accurate, and in almost every case the subjects showed 100% consistent responding to one stimulus in each test pair.

The subject with learning disabilities, MM showed a very different pattern of responding. In phase 1 his transitive inference performance was very stable, showing perfect performance on the $B > D$ inference test. Following the introduction of the $E > A$ inconsistent relation, this performance was seriously disrupted, and on the final test session MM displayed chance level performance on all test pairs and four of the five baseline pairs.

The normally developing young children showed a pattern of responding very similar to that of the normally able adults. The subjects that progressed to phase 2 of the study, generally displayed accurate performance on the baseline pairs and consistent choice on the test pairs, even after the introduction of the inconsistent $E > A$ relation.

This suggests that the normally able adults and children had some way of integrating the inconsistent relation into the previously established series that allowed them to maintain responding. MM did not appear to be able to do this. In this respect the performance shown by MM is very similar to that typically displayed by non-human subjects on this task. Non-human subjects typically display chance level performance on all baseline and test pairs after the introduction of the inconsistent relation, or, in some cases, cease responding altogether. This was what was shown in the experiment by Siemann and Delius

(1994), the human subjects in all groups were able to maintain responding, although response latencies increased with increasing departure from the linear structure. The pigeon subjects showed increasing error rates with increasing departure from the linear structure, and two of the three subjects in group G3, trained on a circular structure ceased responding altogether.

Siemann and Delius gave their human subjects questionnaires and card sorting tasks after the main pair discrimination tasks. The subjects from the group with a linear structure reported that they dealt with the task by memorising the series order and referring back to it. However the group that was trained on a circular structure reported memorising the stimuli pairwise using verbal markers.

Something similar may have happened in this experiment. When the stimuli formed a linear series in phase 1, subjects in each group were able to respond accurately on all pairs. When the $E > A$ relation made the series circular in phase 2, the normally able adults and children may have been able to bring in an alternative strategy, e.g. using verbal markers to remember individual comparisons, to override the now ineffective linear series. Subjects like MM and the non-human subjects may not be able to explicitly utilise an alternative strategy in the same way. If MM was still trying to respond according to a linear strategy, this would explain his chance level performance in phase 2.

Gillan (1981) had suggested that there were two conditions necessary for correct responses on a transitive inference task; accurate performance on the adjacent training pairs (also suggested by Bryant and Trabasso, 1971), and the stimuli that form the series must be ordered on some uni-dimensional scale. Gillan proposed these conditions on the basis of disrupted performance by a chimpanzee subject. The ability of human subjects to continue responding even on a non-linear series might suggest that the second condition is not entirely true. However, it might be more accurate to say that for a subject to respond transitively, the stimuli must form a linear series. The subjects trained on a linear series by Siemann and Delius (1994) used a transitive strategy for responding. While the subjects trained on the circular series were able to continue responding, the strategy they used was not a transitive one.

What this study may show then, is that both human and non-human subjects are able to make transitive inferences. However, when the baseline relations alter so that the series no longer makes transitive sense, normally able humans are able to

respond because they can utilise a non-transitive strategy. This may just be a result of these subjects' greater experience of dealing with abstract or arbitrary relations.

The use of some sort of "strategy" to monitor responding is likely to be a normal part of dealing with these relations. This need not imply that it is *consciously* employed. In this experiment the value attached to each stimulus within a pair was constant. For example, given $B > C$, B is always correct and C is always incorrect. The training in phase 2 removed the end points within the series, making it difficult to ascertain values for non-adjacent pairs, but reinforcement had been given on each adjacent pair comparison, fixing the value for that pair. In a natural setting, the contingencies are unlikely to be arranged as neatly. Instead, situations are likely where one stimulus is correct in one context, while the other would be correct in a different context. The pair comparisons in the transitive inference task (e.g. B compared to C) constitute simple discriminations; three-term contingencies (see Figure 5:2). However, these discriminations will often come under the control of another element, e.g. context 1 or context 2, effectively becoming conditional discriminations; four-term contingencies (see Figures 5:3 and 5:4) such as those described in stimulus equivalence research. Normally able adults generally handle these conditional discriminations with ease, developing complex networks of relations, such as emergent equivalence relations. Indeed, these relations can themselves come under contextual control (see Chapter 11, Experiment 6, for fuller discussion).

The switch from a linear series in phase 1 to a circular series in phase 2 may have mimicked the establishment of contextual control of simple discriminations. As normally able adults are likely to have most experience with relations that change and develop across contexts they would most likely to be able to maintain responding on a series that no longer makes "sense". The establishment of four-term conditional relations has been demonstrated in much younger children than those tested in this experiment (Lipkens, Hayes and Hayes, 1993). It is not surprising therefore that the children in this experiment were also able to maintain consistent responding if the procedure is mimicking the establishment of conditional relations.

The disruption of MM's performance is clearly linked to the introduction of the $E > A$ relations in phase 2. If switching from a linear to a circular series in phase 2 does mimic the establishment of contextual control this may provide some explanation of MM's difficulties. While MM might be able to cope with explicitly

trained conditional discriminations, he may have more difficulty establishing these relations spontaneously himself. Thus, he may have continued to try and respond on the basis of simple discriminations even after the introduction of $E > A$. However, training this relation would have equalised the value for all the stimuli, making it impossible to respond on this basis. This pattern of responding would be similar to the perseveration errors shown by some adults with learning disabilities during stimulus equivalence training. For example, a subject might be taught to select B1 in the presence of A1, and may then continue to respond to B1 once the sample has become A2. This is why an errorless learning procedure was used with the adults with learning disabilities in Experiments 1, 5 and 6 - to minimise the likelihood of these errors.

MM was able to manipulate a logical sequence of simple discriminations, which is necessary for transitive inference performance. This was also demonstrated by the normally able adults and normally developing young children. However, when the series was made circular, effectively putting the pair comparisons under contextual control, the normally able adults and normally developing young children were able to alter their performance to cope with this new task in a way that MM was not. This may suggest an area of particular difficulty for MM.

One result which requires some explanation is the apparent consistent choice shown by the normally able adults on test pair $B > D$ in phase 2. As there was no logical "correct" stimulus in any of the phase 2 test pairs, it would have been expected that roughly half the subjects would select stimulus B and the other half would select stimulus D. This pattern was generally shown on the other two stimulus pairs. However, 9 out of 10 normally able adult subjects consistently selected B over D in the test probes in phase 2, and the other subject showed only a slight preference for D over B. It is likely that this is a direct result of $B > D$ being the transitive inference test pair in phase 1. All the subjects had to select B in preference to D in order to progress to phase 2. Possibly the act of responding to $B > D$, even in the absence of reinforcement, was sufficient to add this pair to the trained baseline relations. As performance on these relations was maintained in phase 2, this may have resulted in the seeming preference for $B > D$. The subject who showed some slight preference for D over B was also the subject who had shown some disruption to the baseline relations in phase 2.

The normally developing young children did not appear to demonstrate this preference for B over D in phase 2; three subjects selected B over D, while two

selected D over B. Possibly the number of subjects involved is too small to show any clear preference. What seems a more likely explanation is that, while training and testing for the adult subjects took place in one session lasting just over one hour, the training sessions for the children were much shorter and took place over an extended period of time, and thus it is more likely then that the adult subjects would remember responding to B in preference to D and still have this relation as part of their baseline.

Table 8:1 Training and Test Stimuli

Stimulus	A	B	C	D	E
Colour	Blue	Red	White	Yellow	Green
Length (cms)	18	15	13	10	8

Table 8:2 Results - Normally able adults, Training and Test Trials, Phase 1

Subject s										
Trial	1	2	3	4	5	6	7	8	9	10
In Order										
A>B	80%	80%	75%	75%	75%	75%	75%	75%	67%	75%
B>C	75%	75%	100%	66%	67%	100%	75%	75%	67%	75%
C>D	100%	100%	100%	75%	100%	75%	100%	100%	100%	100%
D>E	100%	100%	100%	100%	100%	100%	100%	100%	100%	75%
Random Sequence										
A>B	100%	100%	100%	100%	80%	100%	100%	100%	60%	100%
B>C	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
C>D	100%	100%	100%	80%	60%	100%	100%	100%	80%	100%
D>E	100%	100%	100%	100%	80%	100%	100%	100%	100%	100%
Test Probes										
A>B	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
B>C	90%	100%	100%	100%	100%	100%	100%	100%	100%	100%
C>D	100%	80%	100%	100%	100%	100%	100%	100%	100%	100%
D>E	100%	100%	100%	100%	100%	100%	100%	100%	0%	100%
B>D	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Table 8:3 Results - Normally able adults, Training and Test Trials, Phase 2

Subjects										
Trial	1	2	3	4	5	6	7	8	9	10
In Order										
A>B		100%	100%	100%	100%	100%	100%	100%	100%	100%
B>C		100%	100%	100%	100%	100%	100%	100%	100%	100%
C>D		100%	100%	100%	100%	100%	100%	100%	100%	100%
D>E		100%	100%	100%	100%	100%	100%	100%	75%	100%
E>A		75%	75%	75%	75%	75%	75%	75%	75%	75%
Random Sequence										
A>B		100%	100%	90%	100%	100%	90%	80%	100%	100%
B>C		100%	100%	100%	90%	100%	100%	80%	100%	100%
C>D		100%	100%	100%	100%	100%	90%	90%	100%	100%
D>E		100%	100%	90%	100%	90%	90%	70%	100%	100%
E>A		100%	100%	100%	100%	100%	100%	90%	100%	80%
Test Probes										
A>B		100%	100%	100%	100%	100%	100%	100%	100%	100%
B>C		100%	100%	100%	100%	100%	100%	100%	100%	100%
C>D		100%	100%	100%	100%	100%	100%	30%	100%	100%
D>E		100%	100%	100%	100%	100%	100%	70%	100%	10%
E>A		100%	90%	100%	100%	100%	100%	100%	100%	90%
B>D	B	100%	100%	100%	90%	80%	100%	30%	100%	100%
	D	0%	0%	0%	10%	20%	0%	70%	0%	0%
B>E	B	0%	100%	0%	80%	90%	0%	40%	100%	0%
	E	100%	0%	100%	20%	10%	100%	60%	0%	100%
A>D	A	0%	100%	100%	20%	100%	100%	20%	100%	100%
	D	100%	0%	0%	80%	0%	0%	80%	0%	0%

Figure 8:1 Results from test pairs, Normally able adult subjects, phase 2

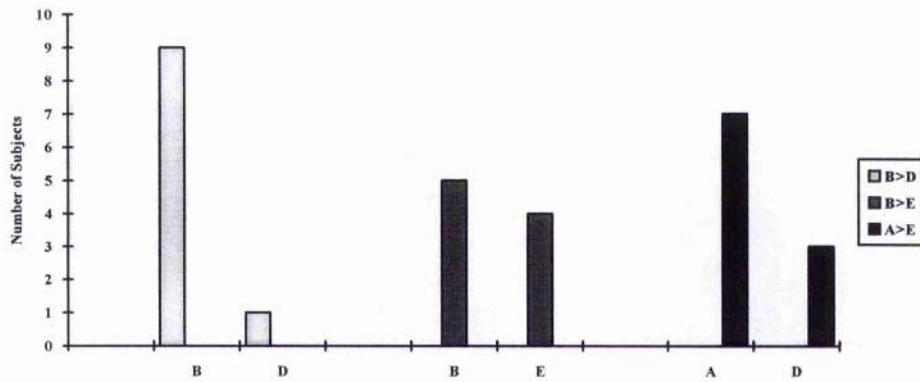


Table 8:4 Adult with learning Disability, Subject MM, Training and Test Results, Phase 1

Trial Type	A>B	B>C	C>D	D>E	B>D
In Order	100	100	83	100	
Random	100	100	80	100	
Random	100	100	78	100	
Random	100	80	100	100	
Test Probes	100	83	100	100	100

Fig. 8:2 Test Probe Results, Subject MM, Phase 1

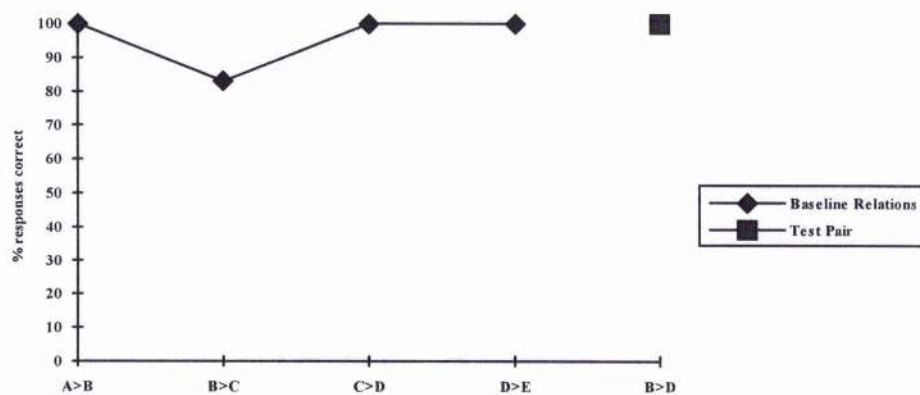


Table 8:5 Adult with Learning Disabilities, Subject MM, Phase 2 Results

	A>B	B>C	C>D	D>E	E>A	B > D		B > E		A > D	
						B	D	B	E	A	D
In Order	100%	75%	75%	100%	75%						
Random	100%	80%	83%	60%	75%						
Random	75%	40%	80%	75%	75%						
Random	38%	30%	50%	50%	67%						
Test Trials	40%	40%	80%	60%	60%	60%	40%	60%	40%	60%	40%

Fig 8:3 Test Probe Results, Subject MM, Phase 2

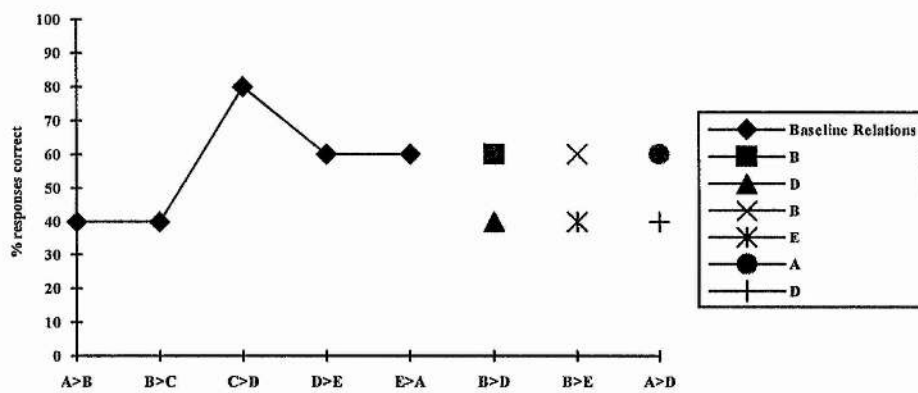


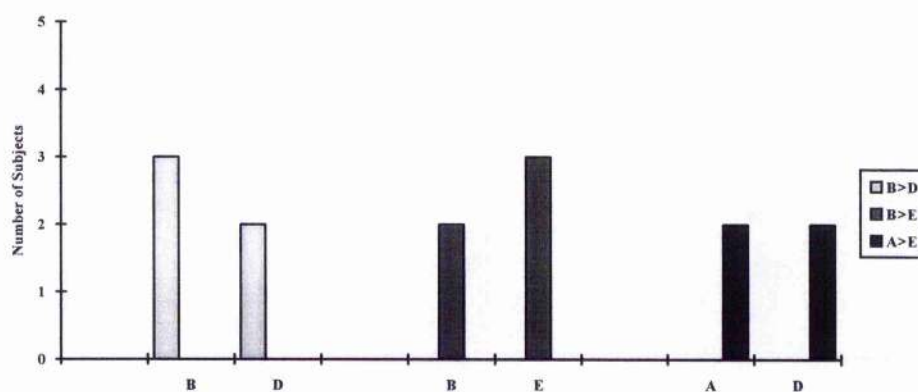
Table 8:6 Test Probe Results, Normally Developing Young Children, Phase 1

Test Pairs	A>B	B>C	C>D	D>E	B>D
Subject 1	100%	100%	0%	40%	100%
Subject 2	100%	80%	80%	80%	80%
Subject 3	100%	100%	80%	100%	80%
Subject 4	0%	80%	40%	100%	20%
Subject 5	60%	20%	60%	80%	80%
Subject 6	100%	100%	100%	80%	100%
Subject 7	100%	100%	100%	100%	80%
Subject 8	100%	100%	100%	100%	100%

Table 8:7 Test Probe Results, Normally Developing Young Children, Phase 2

						B > D		B > E		A > D	
	A>B	B>C	C>D	D>E	E>A	B	D	B	E	A	D
2	100%	100%	100%	83%	100%	75%	25%	92%	8%	25%	75%
3	100%	92%	67%	75%	58%	92%	8%	92%	8%	100%	0%
6	100%	100%	100%	83%	100%	33%	67%	33%	67%	42%	58%
7	92%	100%	100%	100%	92%	25%	75%	17%	83%	33%	67%
8	100%	100%	100%	100%	100%	92%	8%	0%	100%	100%	0%

Figure 8:4 Results from test pairs, Normally Developing Young Children, phase 2



CHAPTER 9

Experiment 4:

Disruption of Stimulus Equivalence Performance

Within Psychology it has been quite common to examine disrupted performance as a way of learning more about normal performance. For instance, much has been learned about language and perception from subjects who have suffered specific brain damage, possibly through stroke or traumatic injury. Alternatively, some studies have tested subjects whose abilities are unimpaired but have examined performance by using procedures designed to test the limits of normal responding. One way of doing this is to disrupt what appears to be normal processing of information. For instance, a Stroop task might present the word RED, very briefly, in either blue or red ink. Subjects would be asked to say what colour the word was presented in. Typically, where the word and the colour are incongruent subjects will have more difficulty perceiving the colour because the actual printed word produces interfering information.

The point of experiments such as these is that it can be hard to perceive the processes operating in normal performance. By studying subjects whose performance is likely to be affected (e.g. as a result of learning disabilities or injury), or by selectively interfering with aspects of "normal" performance, it is possible to learn more about the basis of such skills. This is necessary for more than just academic interest. An inability to perform certain tasks can cause huge problems for an individual, whether the problem is an inability to recognise faces (prosopagnosia) or problems dealing with relational information (such as that seen in transitive inference and stimulus equivalence tasks). It can be very difficult to plan an appropriate intervention to deal with these problems when it is not known what the problem actually is.

Quite a lot is known about performance on standard stimulus equivalence tasks. For instance, it has been shown that stimulus equivalence can help to build complex networks of relational information (e.g. Sidman, Kirk & Willson-Morris, 1985). Less investigation has taken place into what will happen when the relations presented are not all consistent. It is assumed that normally able adults will be able to cope with this - after all, the relational information encountered in real life is rarely as neat as that in an equivalence test. However, inconsistent relations might prove more of a problem for adults with learning disabilities or normally

developing young children whose equivalence performance may not be as well established or adaptable. Selective disruption of one aspect of equivalence relations might give some indication of how inconsistent relations can be processed and integrated into existing relations.

Experiment 3 examined some of the conditions necessary for transitive inference responding by first establishing transitive inference performance, and then selectively altering one aspect of the baseline relations and seeing what effect this had on transitive inference performance. Following a disruption of this kind with a chimpanzee subject Gillan (1981) had suggested that appropriate transitive inference performance was based on accurate performance on the adjacent baseline relations, and that the stimuli making up the series must be ordered on some uni-dimensional scale. The results from Siemann and Delius (1994), and experiment 4, suggested that normally able subjects could in fact continue to respond consistently on a circular series but they did not use a transitive strategy when they did so.

One reason that normally able humans, both adults and children, may be able to continue responding is that they may be used to tasks with abstract or arbitrary relations between stimuli. The transitive inference task used was based on a greater than/less than relation between the stimuli ($A > B$, $B > D$ etc.). When the series was made circular it destroyed the basis for the relation between the stimuli. In stimulus equivalence however, the relations between the baseline relations are entirely arbitrary anyway. According to Hayes (1991), one of the defining characteristics of stimulus equivalence is that it is an example of arbitrarily applicable relational responding. In this case, would a disruption of the baseline relations cause similar disruption to equivalence responding in subjects who had previously displayed stable equivalence relations? The aim of this experiment was to establish equivalence responding and then selectively disrupt one of the derived relations, and investigate if this would affect the previously displayed equivalence relations.

Stimulus equivalence requires the demonstration of the properties of reflexivity, symmetry and transitivity. All of these properties must be present if equivalence is demonstrated. For instance, if a subject is taught the conditional relations A-B and B-C, then stimulus equivalence requires the demonstration of the emergent C-A relation. If this relation is demonstrated, then it can be assumed that the requisite symmetrical and transitive relations have also been derived. In this experiment, two baseline conditional discriminations were trained, e.g. A-B and B-C. The subjects were then tested for the emergence of the C-A relation. Provided this was

demonstrated, the subject received training on the A-C transitive relation that was inconsistent with the relations that had previously been established. This transitivity relation had never been explicitly tested but can be assumed to be present as the equivalence relations had been demonstrated. Performance on the C-A equivalence relation was then reassessed to see if the inconsistencies between the derived transitivity relation, and the taught transitivity relation caused any disruption to equivalence performance.

This experiment was carried out with three groups of subjects, normally able adults, an adult with learning disabilities, and normally developing young children. The normally able adults were expected to be most likely to integrate the inconsistent relation and continue to respond consistently as this group of subjects demonstrates equivalence relations most reliably, and has most experience with arbitrary relations. In Experiment 3, the adult with learning disabilities had shown substantial disruption of performance following the introduction of the inconsistent relation. However, in this experiment, for the subject to receive inconsistent transitivity training, the subject had first to demonstrate the arbitrary equivalence relations, and thus was clearly capable of responding on an arbitrary task, at least while all the taught relations were consistent.

It was also not entirely clear how well the normally developing children would cope with this task. Arbitrary equivalence relations have been demonstrated with children aged two and even slightly younger. Therefore young children clearly manipulate arbitrary relations. However, these children's performance on equivalence tasks is unlikely to be as robust and well established as that of normally able adults, and thus may be more susceptible to disruption by the inconsistent transitivity relation.

If the subjects demonstrated equivalence relations in phase 1 and then received inconsistent transitivity training in phase 2, there seemed to be three likely outcomes when tested for equivalence once more. The subjects could demonstrate equivalence relations based on the training received in phase 1, effectively ignoring the inconsistent transitivity relations; they could demonstrate new equivalence relations based on these transitivity relations; or the inconsistent relations could cause general disruption of performance, with no clear equivalence relations being demonstrated.

Method

Subjects

Three different groups of subjects were tested:

1. Normally able adults.

Ten undergraduate students were recruited by means of a sign-up sheet in the Psychology department. Seven of these subjects had also taken part in experiment 4. As before, participation was limited to first year psychology students or non-psychology students to ensure the subjects were not familiar with the area of study. The subjects were paid for their participation.

2. Adult with learning disabilities.

The subject was JD, a female aged 71. JD had previously taken part in another experiment on equivalence (experiment 1). JD was paid for her participation, receiving 50p at the end of each session.

3. Normally developing young children.

These subjects were nine pre-school children aged between 3:2 and 4:10 at the time each began testing (mean age 3:11; years:months). (see Table 9:1) All the children attended a pre-school playgroup attached to the School of Psychology.

Stimuli

The stimuli used for training and testing were upper and lower case Greek letters and their printed names (see Table 9:2). Two sets of stimuli were used, the letters gamma and delta were used with the normally able adults and the young children. As JD had previously received training with these stimuli, she was trained using the letters xi and phi.

An errorless training procedure was used with subject JD. This meant that six cards were prepared for each individual stimulus at different levels of intensity, from pale grey on white to full black on white. For the normally able adults and young children one card was used for each stimulus, as all the stimuli always appeared at full intensity black on white.

Additional stimuli used only with the young children were a container with 40 wooden beads and a small glass jar.

Design

The basic design of the experiment was the same for all subjects and was carried out in two phases.

Phase 1.

The aim of phase 1 was to establish stable equivalence relations. The subjects were taught four conditional discriminations, e.g. A1-B1, A2-B2, B1-C1 and B2-C2. They were then tested for the emergence of equivalence relations e.g. C1-A1, C2-A2. Provided they demonstrated equivalence relations the subjects progressed to phase 2 of the study.

Phase 2.

In phase 2, two further conditional relations were taught. These new relations were inconsistent with the transitivity relations assumed to have been derived in phase 1. The demonstration of equivalence in phase 1 would have required the emergence of the relations A1-C1 and A2-C2. In phase 2 the subjects were taught the conditional relations A1-C2, A2-C1. The subjects were then tested once more on the c-A equivalence relations to see if these new transitivity relations had caused any changes to the previous performance.

Procedure

The procedure used varied slightly for each group of subjects.

1. Normally able adults.

In phase 1 these subjects were taught the conditional relations A1-B1, A2-B2 and A1-C1 A2-C2, and tested on the equivalence relation C-B. These discriminations were trained using a trial and error procedure to a criterion of five consecutive correct responses on each discrimination. Initially the subject was taught the two A-B conditional discriminations. When they had reached criterion on these they were tested on the A-B mix, where either stimulus A1 or A2 could function as the sample stimulus, each being tested five times in a block of ten trials. Criterion for passing the mix was set at 90% correct. The A-C discriminations were then trained and the A-C mix tested in the same way. The subjects were then tested on the C-B equivalence relations. Once again criterion was set at 90% correct responding. Provided the subject demonstrated appropriate equivalence relations they progressed to phase 2 of the study. No feedback was given on either the tests of the A-B or A-C mixes or the equivalence test. On all trials the left/right position of the comparison stimuli varied randomly.

In phase 2 the subjects were taught the conditional discriminations B1-C2, B2-C1, again to a criterion of five consecutive correct responses. The subjects were then tested once more on the C-B equivalence relations.

2. Adult with learning disabilities.

In phase 1 JD was taught the conditional discriminations A1-B1, A2-B2, B1-C1, B2-C2 and tested on the equivalence relation C-A. These discriminations were taught using an errorless learning procedure. On each trial a sample stimulus was presented (e.g. A1) followed by two comparison stimuli, e.g. B1 (S+) and B2 (S-). On each trial the sample and correct comparison (S+) stimuli appeared at full intensity black on white. The incorrect comparison (S-) initially appeared as pale grey on white. As correct responding to S+ was established, the incorrect comparison slowly darkened in colour until it too was at full intensity black on white. There were six different levels of intensity. Three consecutive correct responses were required at each level before the intensity was increased. An incorrect response at any level resulted in the intensity of the S- stimulus immediately being reduced by one level until performance was stable once more. Once both the correct and incorrect comparison stimuli were at full intensity six consecutive correct responses were required before the subject was said to have reached criterion on that discrimination.

Using this procedure JD was taught the relations A1-B1 and A2-B2. She was then tested on the A-B mix. As with the normally able subjects, criterion was set at 90% correct responding. The errorless learning procedure was used to teach the relations B1-C1 and B2-C2 and the B-C mix was tested. Once the A-B and B-C mixes were both at criterion JD was given the C-A equivalence test. Once again criterion was set at 90% correct.

In phase 2, the errorless learning procedure was used to teach the relations A1-C2, A2-C1. Once JD had reached criterion on both these relations she was tested once more on the C-A equivalence relations. After testing C-A equivalence, JD was tested once more on the previously stable A-B and B-C mixes.

3. Normally developing young children.

In phase 1 the subjects were taught the conditional discriminations A1-B1, A2-B2 and B1-C1, B2-C2, and tested for the emergence of the C-A equivalence relations. These discriminations were taught using a trial and error procedure. Initially the

conditional discriminations A1-B1 and A2-B2 were trained. For each conditional discrimination the criterion was 15 correct responses in a block of 16 trials. All correct responses were reinforced. When the children reached criterion on these discriminations individually, they were tested on the A-B mix. The A-B mix was presented in blocks of 10 trials and criterion was set at 9/10 trials correct. The children continued to receive reinforcement on the first block of 10 trials. If they reached criterion on that block, on the next block they were tested for the maintenance of the A-B relations in the absence of reinforcement.

Provided the children maintained criterion performance on this task, training was then given on the B-C relations B1-C1 and B2-C2. These were trained in the same way as the A-B relations and criterion was again 15/16 trials correct. Once criterion was achieved on these relations the B-C mix was tested. As before, reinforcement was provided on the first block of 10 trials and, as long as criterion was reached, the subjects were tested for maintenance of this performance in the absence of reinforcement.

The subjects were then tested on the A-B and B-C mixes in a single session. This session consisted of 4 blocks of 10 trials. Blocks 1 and 2 tested the A-B mix and blocks 3 and 4 tested the B-C mix. In the first session of this type, reinforcement was provided on blocks 1 and 3 with no reinforcement on blocks 2 and 4. Provided the subjects reached criterion on both discriminations, the two discriminations were again tested in a single session, but this time no reinforcement occurred on blocks 1 and 3, reinforcement was provided on blocks 2 and 4. If the subjects maintained criterion responding on this session, they received C-A equivalence testing on the following session.

In the session testing for the emergence of C-A equivalence, the subjects were presented with 2 blocks of 4 trials, one of the A-B relations and one of the B-C relations, both reinforced. They were then presented with one block of 10 C-A trials where either C1 or C2 functioned as sample on each trial, with A1 and A2 as the comparisons. No reinforcement was provided on this block. The criterion on this task was again 9/10 trials correct, and the subjects progressed to phase 2 of the study only if they achieved criterion on the C-A equivalence trials.

In phase 2 a similar training procedure was used to teach the relations A1-C2 and A2-C1. The relations were first taught individually to a criterion of 15/16 correct responses. They were then presented in the A-C mix. The mix was presented in 2 blocks of 10 trials. The first block was reinforced and the second block tested for the maintenance of A-C performance in the absence of reinforcement. provided the

subjects reached criterion on this task they were given the C-A equivalence test for the second time on the next session.

In this session the subjects were given 3 blocks of 4 trials, one block of A-B relation, one of B-C relations and one of A-C relations. Reinforcement was provided on all 3 blocks. The subjects were then given one block of 10 trials testing the C-A equivalence relations. No reinforcement was provided on this block.

Some additions were made to the procedure with regard to feedback and reinforcement. In both phases 1 and 2, before all training sessions, a container with 40 wooden beads and a small glass jar were set on the table in front of the subject. Before the first training session the subject was told that they were going to play a game in which they could earn beads, and that every time they made a correct response they would earn one bead. When all the beads in the tray had been earned (put in the glass jar), the subject received a pre-selected back up reinforcer (a small picture). If, during a training session, the subject earned all the beads, the beads were exchanged for the pre-selected picture, the beads were emptied back into the container, and a new picture was selected. The training session then continued as before. Before test and maintenance sessions the subject was shown the glass jar with the beads earned during the previous session. The jar and tray of beads were then removed from the table and the subject was told, "Now we are going to play the game without me telling you if you are right or wrong. You will also not get any beads while we are playing, but I will keep a record of how many beads you earn and you will be given your beads at the end of the game. Do your best". At the end of the test session, the subject was given one bead for every trial completed.

Results

The results from each group of subjects will be presented separately. Responding on the equivalence test in phase 2 is discussed in terms of whether the subjects responded consistently according to the relations established in phase 1, e.g. C1-B1, C2-B2, or according to the relations established in phase 2, e.g. C1-B2, C2-B1, or whether the subjects failed to respond consistently on this test.

1. Normally able adults

All the subjects learned the A-B and A-C discriminations in phase 1 easily, and performed accurately on the A-B and A-C mixes (see Table 9:3). When tested, 9 of the 10 subjects immediately demonstrated appropriate equivalence relations.

Subject 9 initially selected stimulus B2 given sample C1 and stimulus B1 given sample C2. However, after six trials of equivalence testing, this subject said he thought he had made a mistake and wanted to change his response. He then began to select B1 given C1 and B2 given C2. The subject was then given another block of 10 trials and demonstrated appropriate equivalence relations throughout this second block. All 10 subjects then progressed to phase 2.

In phase 2 all subjects made a "mistake" on the first trial of B1-C2. This was expected as this relation is inconsistent with the relations derived in phase 1. However all the subjects then reached criterion on B1-C2 and all immediately demonstrated accurate performance on B2-C1. When retested on the C-B equivalence relations, all subjects responded consistently on this task to one or other of the sets of equivalence relations that had been established. Three subjects (subjects 1, 6 and 9) responded according to the relations established in phase 1, and the other 7 responded according to the equivalence relations established in phase 2.

2. Adult with learning disabilities

In phase 1 JD reached criterion on both A1-B1 and A2-B2 the first time each discrimination was presented. She also displayed perfect performance on the A-B mix (see Fig. 9:1). JD reached criterion on both B1-C1 and B2-C2 the first time each was presented, but performance on the B-C mix was below criterion (70%). There was then an unavoidable break in testing for five days. Following that the A-B and B-C relations were retrained and performance on both the A-B and B-C mixes was perfect. JD was then tested for the emergence of C-A equivalence and performance on this relation was at criterion.

JD then received training on the inconsistent A-C relation. She reached criterion on both A1-C2 and A2-C1 the first time each was presented. JD was then tested on the C-A equivalence test once more but demonstrated chance level responding (50%; see Fig 9:2). JD was tested on the A-B and B-C mixes as established in phase 1, but showed considerable disruption on both of these; A-B mix - 30%, B-C mix - 60%.

3. Normally developing young children

Nine subjects began training in phase 1. During the course of phase 1, four subjects were dropped from further participation in the study. Subjects 1 and 5 were dropped as they were only available for training and testing infrequently. Subject 7

was dropped after 12 sessions as she consistently failed the A-B mix, despite repeated training. Subject 8 was dropped after 6 sessions as she was consistently unable to reach criterion on B1-C1. The other five subjects required between 5 and 17 sessions to complete phase 1 training and testing. Subjects 2, 4, 6 and 9 all demonstrated appropriate C-A equivalence relations and progressed to phase 2 of the study (see Fig. 9:3). Subject 3 was tested twice for the emergence on C-A equivalence, but on both occasions performance was at chance level. This was despite an extensive history of training on the baseline discriminations, and accurate performance on both the A-B and B-C relations. As a result this subject was dropped from further participation in the study.

In phase 2, the four remaining subjects all learned the A1-C2 and A2-C1 relations. Subjects 2, 4 and 6 all showed appropriate performance on the A-C mix. Subject 9 was slightly below criterion on the A-C mix (80%; see Table 9:4), but as the time available for testing was limited, and performance had been stable on the second half of the test these discriminations were not retrained. Subjects 2, 4 and 9 received the final C-A equivalence test. Subject 6 was unavailable for this final test. Before the final C-A equivalence test the subjects were given a brief review of the relations that had been trained in phases 1 and 2 of the study. Four reinforced trials were given on each of the A-B, B-C and A-C relations. All the subjects showed rather variable performance on these relations. All three subjects showed rather unstable performance on the C-A equivalence relations. Subjects 2 and 4 responded to the new equivalence relations 60% of the time and subject 9 responded to the new equivalence relations 70% of the time (see Fig. 9:4). For subjects 4 and 9 the "errors" (responding to the original equivalence relations) were demonstrated randomly throughout the test performance. Subject 2 initially responded consistently to the equivalence class established in phase 1, but after four trials switched, and responded consistently to the equivalence class established in phase 2.

Discussion

In this experiment, all groups of subjects demonstrated the emergence of untaught equivalence relations at the end of phase 1. When the inconsistent transitivity relations were trained in phase 2, the ability of the subjects to deal with these relations differed substantially.

The normally able adults seemed to be able to assimilate this information. After a maximum of two "errors" at the beginning of the B1-C2 training, the subjects were able to respond consistently in line with the new contingencies, even on the first trial of the B2-C1 relations. When tested on the C-B equivalence relations for a second time these subjects responded consistently to the equivalence relations established in phase 1 (three subjects), or the equivalence relations established in phase 2 (seven subjects). None of the subjects showed intermediate responding. In contrast, the subject with learning disabilities, JD, showed substantial disruption of performance following the introduction of the inconsistent relation. JD's performance on the relations in phase 1 was highly accurate prior to testing, and her performance on the C-A equivalence test was at criterion. JD showed no difficulty in learning the inconsistent A-C transitivity relation, but when tested on the C-A relations once more, was unable to respond consistently to the relations established in phase 1 or phase 2, instead showing chance level responding. This suggests that although JD was able to learn both the consistent and inconsistent relations, she was unable to integrate the conflicting information. Alternatively, consistent responding in phase 2 may require the subject to be able to ignore one set of information, basing responding only on information from phase 1 or phase 2. JD's performance suggests that she was not able to do this either. After being tested on the C-A equivalence relations for a second time, JD was tested once more on the A-B and B-C relations from phase 1. These relations both showed considerable disruption from their previously stable performance. This meant that not only was JD unable to respond consistently on the equivalence relations, but that the conflicting information was causing disruption to the other relations, suggesting that JD had tried to integrate the two sets of information but could not find a way to do this.

If this is true, it would have been interesting to see if the normally able adult subjects were able to maintain performance on the A-B and A-C relations. Three subjects responded consistently to the first set of equivalence relations, which would suggest they would be able to maintain these baseline A-B and A-C relations. In that case, would they then fail to maintain performance on the inconsistent A-C relations? Similarly, seven subjects responded consistently to the equivalence relations established in phase 2. Would they maintain A-C performance while showing disruption to the A-B and A-C relations, or are these subjects able to maintain performance on conflicting items of information and still respond consistently on the equivalence test?

Like subject JD, the normally developing children seemed to have some difficulty integrating the inconsistent items of information. The three subjects who completed phase 2 of the experiment had all shown stable equivalence relations in phase 1, and appeared to have no difficulty learning the inconsistent A-C relations in phase 2. However, performance on the C-A equivalence relation appeared to be rather unstable for all three subjects. This may have been because these were the only subjects to receive a review of the trained relations from phases 1 and 2 immediately prior to equivalence testing. Both the normally able adults and the learning disabled adult could have based performance on either of the sets of trained relations, while ignoring the other set. The children were required to review the trained relations before testing and this may have made them more likely to show inconclusive equivalence performance. However, JD's inconsistent performance suggests that she was trying to integrate the trained relations from both phases, and as three of the normally able subjects responded consistently to the equivalence relations established in phase 1, this suggests that the subjects did not automatically base their responding on the most recently trained relations.

It may just be that the normally able adults will have had much more experience with arbitrary relations and handling seemingly incompatible items of information. Results from the experiment on transitive inference by Siemann and Delius (1994) suggested that when the relational information could no longer be ordered on a single premise, normally able adults are able to switch to an alternative strategy (in this case using verbal markers) to maintain responding. This may be what the normally able adults have done in this experiment. Subject JD is clearly able to coordinate arbitrary relations as she demonstrated equivalence responding, but she may only be able to do this while these relations are consistent. As she is likely to have less experience with purely arbitrary relations, she may not have developed alternative strategies to deal with inconsistencies in the baseline relations.

This would also seem likely for the normally developing children as they also showed inconsistent equivalence relations in phase 2. In fact there is some evidence that subject 2 may have almost developed the ability to deal with inconsistent premises. On the final equivalence test, this subject's overall performance was around chance level (60%). However, this performance was made up of four consistent responses to the phase 1 equivalence relations, followed by six consistent responses to the phase 2 relations. This may mean that the subject was able to maintain consistent responding, but chose to switch classes in the middle of testing. Possibly further testing would have shown stable performance.

(However, Saunders and Saunders, 1994, have warned of the dangers of continued testing until stable performance is demonstrated).

It should be noted however, that there are differences in the procedures used with the different groups, and these might have affected the results. At the very least they make it difficult to compare the results from the different groups with confidence. For instance, the normally developing children were given reviews of all the baseline relations trained, from both phases 1 and 2, before being retested on the equivalence relations. It would also be interesting to know how the introduction of the inconsistent transitivity relation affects performance on the other trained relations.

Another problem was that this experiment only presents data from one adult with learning disabilities. While JD's results are important as she had shown such stable equivalence performance previously, it would be preferable to have results from a larger number of adults with learning disabilities.

In order to address these issues Experiment 4 was repeated, with a tighter training and test procedure, with two additional groups of normally able adults and adults with learning disabilities. The details of this replication are contained in Chapter 10: Experiment 5.

Table 9:1 Chronological ages of normally developing young children

Subject Number	Age (years:months)	Subject Number	Age (years:months)
1	4:4	6	3:10
2	4:7	7	3:7
3	3:8	8	3:2
4	3:7	9	4:10
5	4:0		

Table 9:2 Training and Test Stimuli

Stimulus Class	A	B	C
----------------	---	---	---

Normally able adults and young children

1	Γ	γ	gamma
2	Δ	δ	delta

Adult with learning disabilities

1	Ξ	ξ	xi
2	Φ	ϕ	phi

Table 9:3 Results - Normally Able Adults, Training and Test Trials, Phases 1&2

Subjects										
	1	2	3	4	5	6	7	8	9	10

Phase 1

A1-B1	83%	100%	100%	83%	88%	100%	100%	100%	86%	100%
A2-B2	100%	83%	80%	100%	100%	100%	100%	100%	100%	100%
A-B mix	90%	100%	100%	100%	100%	100%	100%	100%	100%	100%
A1-C1	100%	100%	100%	100%	100%	100%	83%	100%	100%	100%
A2-C2	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
A-C mix	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
C-B equiv.	90%	100%	100%	100%	100%	100%	100%	100%	60% 100%	100%

Phase 2

B1-C2	83%	83%	83%	71%	83%	83%	83%	83%	83%	83%
C2-B1	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
C-B equiv.	20%	100%	100%	100%	100%	100%	20%	100%	20%	100%
Equiv. Class	1st	2nd	2nd	2nd	2nd	2nd	1st	2nd	1st	2nd

The notation 1st or 2nd in phase 2 refers to whether the subjects responded on the C-B equivalence test in this phase, according to the relations established in the first phase of the experiment, or the relations established in the second phase of the experiment.

Fig 9:1 Subject JD, Training and Test Results, Phase 1

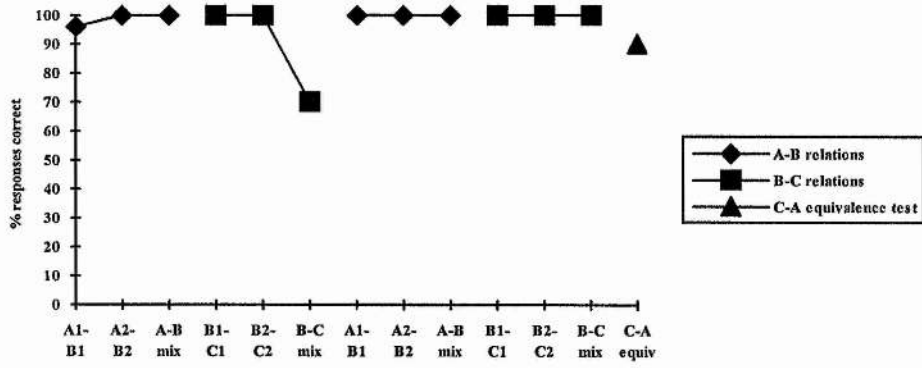


Fig 9:2 Subject JD, Training and Test Results, Phase 2

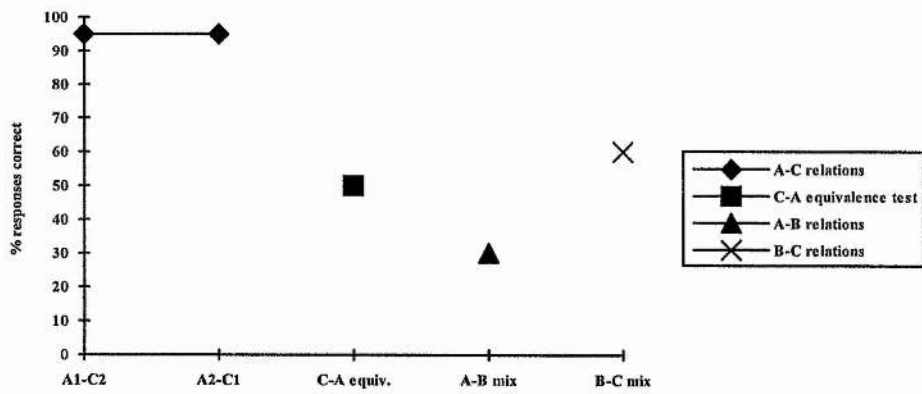


Fig. 9:3 Normally Developing Young Children
Results from Subjects Completing Phase 1

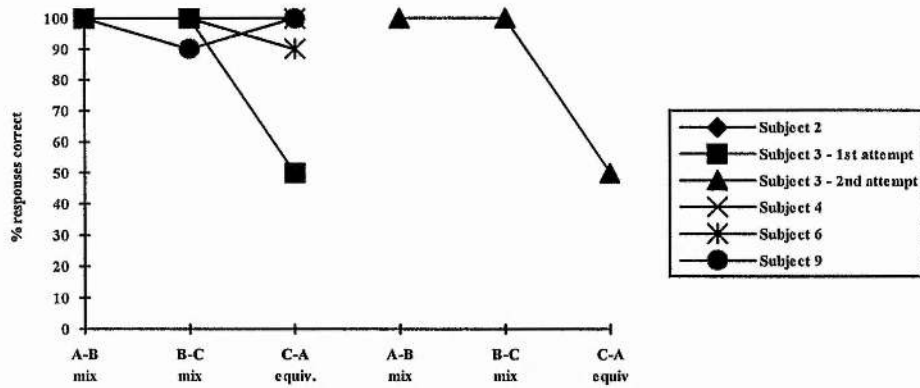
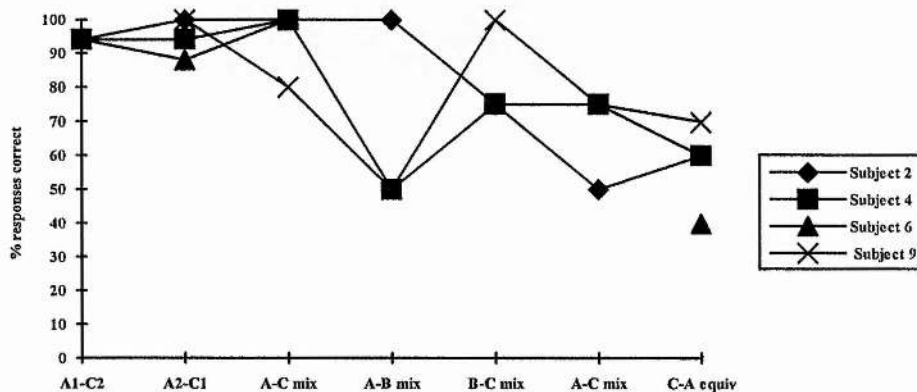


Table 9:4 Normally Developing Young Children
Results from Phase 2

Subject				Final	Test	Session	
	A1-C2	A2-C1	A-C mix	A-B mix	B-C mix	A-C mix	C-A equiv.
2	94%	100%	100%	100%	75%	50%	60%
4	94%	94%	100%	50%	75%	75%	60%
6	94%	88%	100%				
9	94%	100%	80%	50%	100%	75%	70%

Fig 9:4 Normally Developing Young Children
Results from Phase 2



CHAPTER 10

Experiment 5:

Replication of Experiment 4 - Disruption of Stimulus Equivalence Performance

This experiment is a replication and extension of the work carried out in Experiment 4. That experiment aimed to establish stable stimulus equivalence performance, and then selectively to disrupt one of the derived relations by training conditional relations incompatible with the stimulus classes already established. The effects of these incompatible relations on the derived equivalence relations was then examined.

This procedure was carried out with groups of normally able adults, normally able young children, and an adult with learning disabilities. The procedure caused substantial disruption to performance for the adult with learning disabilities. The normally able young children showed at least some disruption to performance (the amount of disruption varying for each subject). However, the normally able adults were able to maintain consistent responding to the relations established in either phase 1 or phase 2 of the experiment (see Chapter 9 for fuller discussion).

The normally able adults were the only subjects able to show consistent responding on the equivalence test following the introduction of the inconsistent relations. These subjects may have found a way to integrate the inconsistent pieces of information and thus show consistent responding. Alternatively, the subjects may have "chosen" to ignore one set of the conflicting relations and thus respond consistently by explicitly monitoring their responses.

The normally able children and the adult with learning disabilities did not appear to use either of these strategies. This may have been because they lacked skills or experience in dealing with arbitrary or conflicting relations.

However, differences in the procedures used may make it difficult to compare the results from the three groups of subjects. The aim of the experiment was to investigate the effects of introducing a new relation inconsistent with relations previously trained and derived. This assumes that the subjects will attempt to integrate the two sets of relations. However, only the procedure used with the normally able children *explicitly* required the subjects to attempt this.

In the procedure used with the normally able adults and the adult with learning disabilities, phase 1 trained the baseline conditions necessary for the subjects to derive stimulus equivalence, and subsequently tested for the emergence of the equivalence relations. The inconsistent relation was introduced in phase 2 and equivalence performance was retested but this inconsistent relation was never presented with the baseline relations from phase 1. This might make it easier to “disregard” one set of relations and thus respond consistently *either* to the relations established in phase 1 *or* to the relations from phase 2. This would be one possible explanation for the results shown by the normally able adults.

It seems unlikely that this happened with JD, the adult with learning disabilities, whose training and testing followed the same procedure. JD showed appropriate equivalence responding in phase 1 of the experiment but chance level responding on the equivalence test following the introduction of the inconsistent relations.

The procedure used with the normally able children was more likely to cause the subjects to try and integrate the sets of conflicting relations. The order of training and testing in phase 1 was the same as that used with the other subjects. The subjects were taught the A-B and B-C conditional discriminations and tested for the emergence of the C-A equivalence relations. However, in phase 2, after the introduction of the inconsistent A-C conditional relations the subjects were required to review the A-B and B-C relations from phase 1, as well as the A-C relations from phase 2, before the C-A equivalence relations were retested. This procedure makes it more likely that the subjects would try and integrate the inconsistent relations. As the aim of the experiment was to investigate the effects of training conflicting baseline relations, this was why the review session was introduced to phase 2 of the training.

However, while the normally able children were the only subjects in this experiment *explicitly* required to review all the trained relations before the final equivalence test, it would be unwise to assume that the normally able adults and the adult with learning disabilities did *not* try and integrate the relations. One of the most striking features of stimulus equivalence is the potential for very large and complex equivalence classes to develop. Sidman, Kirk & Willson-Morris (1985) established three six-member equivalence classes (15 directly taught conditional relations, resulting in 60 new untaught relations). Even more complex sets of equivalence relations are thought to develop in everyday life. The study by Lipkens, Hayes & Hayes (1993) documented the demonstration of emergent

equivalence relations in a human child from the age of 16 months 18 days to 27 months 11 days and numerous studies have discussed stimulus equivalence as an example of a powerful *natural* learning process. Indeed, Sidman (1994) has suggested that initially all stimuli enter into equivalence relations and that it is experience which breaks these down into appropriate equivalence classes.

It is possible then, that the normally able adults in experiment 4 were able to maintain consistent responding as a result of experience in dealing with conflicting items of information. These subjects may indeed have tried to integrate the conditional discriminations trained, and were able to respond consistently as a result of overt strategies developed through experience with conflicting items of information.

The aim of this follow-up experiment was to repeat the procedure used with the normally able children with normally able adults and adults with learning disabilities as subjects. It was assumed that the normally able adults in the previous experiment did attempt to integrate the conflicting relations but were able to maintain consistent responding, possibly by developing strategies to “monitor” their responding or by choosing to disregard one set of conflicting relations. If this is true, then programming a review of all the trained baseline relations should not affect performance. However, the normally able adults might have maintained consistent responding by *not* integrating the conflicting relations, possibly by treating phase 1 and phase 2 of the experiment as essentially different tasks. In that case it is likely that the review of the inconsistent training relations in phase 2 would cause disruption to performances on the final equivalence test by requiring the subject to try and integrate the inconsistent relations.

The rationale for repeating this experiment with more adults with learning disabilities is slightly different. It seems likely that JD who originally took part in this experiment did try to integrate the inconsistent relations and this caused considerable disruption to performance on the final equivalence test. JD had shown stable equivalence performance in phase 1 of the experiment but this was disrupted following the introduction of the inconsistent relations in phase 2 of the experiment. As the method of training and testing was the same in phase 2 as phase 1 the most likely explanation for this disruption was that she had tried to integrate the training relations and had not found a way to deal with the inconsistencies. Moreover, JD had very clearly demonstrated equivalence on two previous

occasions (experiment 1 and unpublished data) and so this disruption seems unlikely to be due to unstable performance on equivalence tasks.

However, while the effects of the manipulation on JD's performance seem clear, it is still only data from one subject and the conclusions that can be drawn on the basis of that are limited. It would be preferable to have data from at least three or four more subjects to gain some idea of how typical this performance is. It was decided, therefore, to repeat the procedure with at least four more adults with learning disabilities. While it seems likely that JD did try and integrate the training relations, it cannot be assumed that all subjects given training and testing with the original procedure would do so. Thus, in this replication, the sequence of training and testing used in the original experiment with the normally able young children was used.

Method

Subjects

Two groups of subjects were tested

1. Normally able adults

Four normally able adult subjects took part. Two were male, subjects 1 and 4 and two were female, subjects 2 and 3. All subjects were assistant psychologists or undergraduate volunteers in a clinical psychology department. None of the subjects were familiar with this area of psychology.

2. Adults with learning disabilities

Four adults with learning disabilities took part. Three were male, Subjects 5, 7 and 8 and one was female, Subject 6.

Stimuli

The stimuli used for training and testing were upper and lower case Greek letters and their printed names. Two sets of stimuli were used (see Table 10:1). The letters xi and phi were used with the normally able adults and the letters gamma and delta with the adults with learning disabilities.

For the normally able adults one card was used for each stimulus, all stimuli appearing as full intensity black on white. An errorless learning training procedure was used with the adults with learning disabilities. This meant that six cards were

prepared for each individual stimulus, at different levels of intensity, from pale grey on white to full intensity black on white.

Design

The basic design of the experiment was the same for all subjects and was carried out in two phases.

Phase 1

The aim of phase 1 was to establish stable equivalence performance. The subjects were taught the A-B conditional discriminations A1-B1 and A2-B2, and the B-C conditional discriminations B1-C1 and B2-C2. Performance on these relations was reviewed before testing for the emergence of the C-A equivalence relations C1-A1 and C2-A2. Phase 1 was designed to ensure highly accurate performance on the trained conditional discriminations and criterion for passing the C-A equivalence test was also high (90%) to ensure that derived equivalence performance was stable.

Phase 2

In phase 2, two further conditional relations were taught. These new relations were inconsistent with the A-C transitivity relations assumed to have been derived in phase 1. The demonstration of equivalence in phase 1 would have required the emergence of the relations A1-C1 and A2-C2. In phase 2 the subjects were taught the conditional discriminations A2-C1 and A1-C2. The baseline A-B and B-C relations taught in phase 1 and the inconsistent A-c relations taught in phase 2 were then all reviewed prior to retesting the C-A equivalence relations. This review session was designed to require the subjects to treat the relations trained separately in phases 1 and 2 as part of the same task, and not as separate tasks.

Procedure

The procedure used varied slightly for each group of subjects. In each case, the order of training and testing was designed to maintain performance on the baseline relations at as high a level as possible.

For both groups of subjects, throughout the procedure, the subjects were told before each set of trials whether they would or would not receive reinforcement. The order in which the discriminations were presented varied in a pseudo-random manner so that each discrimination (e.g. A1-B1 or A2-B2) occurred the same

number of times within each block of trials. On all trials the left/right position of the comparison stimuli varied randomly.

1. Normally able adults

In phase 1 these subjects were taught the conditional relations A1-B1, A2-B2, B1-C1 and B2-C2 using a trial and error procedure. Initially the subjects were taught the two A-B conditional discriminations together, with reinforcement, to a criterion of 90% correct. When the subjects reached criterion on the reinforced trials they were then given ten trials without reinforcement to test for maintenance of the relations, criterion again being 90% correct. The subjects were then taught the B-C relations and maintenance was checked in the same way. Performance on the A-B and B-C mixes was reviewed with subjects receiving first six trials with reinforcement and then six trials without reinforcement on each set of relations.

In the session testing for equivalence, maintenance of these relations was reviewed once more with subjects receiving four trials without reinforcement then four trials with reinforcement on each set of relations. This was followed by eight trials on the C-A equivalence relations with no reinforcement available. The aim of this procedure was to check that the trained relations would be maintained in the absence of reinforcement and the brief sets of reinforcement ensured that there was no extinction of the baseline relations prior to testing of the C-A equivalence relations.

In phase 2 the A-C relation were trained, and maintenance tested in the same way. The subjects were taught the conditional discriminations A1-C2 and A2-C1 together, first with reinforcement, then presented without reinforcement as a maintenance check. The amount of training given on the A-C relations in phase 2 was matched to that needed to train the A-B and B-C relations in phase 1. As the A-C relations were incompatible with the relations trained previously, this might have affected the amount of training required to achieve a set criterion on the A-C task. Indeed, it was possible that the subjects might fail to learn the A-C relations at all. This was why training was matched to that required to establish stable A-B and B-C performance in phase 1.

Following training on the A-C relations, the A-B and B-C relations were reviewed, subjects receiving first six trials with reinforcement, then six trials without reinforcement on each set of relations. As in phase 1, in the sessions testing for C-

A equivalence, maintenance of all baseline relations was tested once more. Subjects received four trials without reinforcement followed by four trials with reinforcement on each set of relations. This tested to see if any disruption would be shown to performance on the trained baseline relations while ensuring that any disruption to performance on the final equivalence test was not due to extinction in the absence of reinforcement. Following this review the subjects received eight trials without reinforcement, retesting performance on the C-A equivalence relations.

2. Adults with learning disabilities

The order of training and testing used with these subjects was the same as that used with the normally able adults. The procedure differed in that an errorless learning training procedure was used to teach the A1-B1, A2-B2, B1-C1, B2-C2, A1-C2 and A2-C1 conditional discriminations. On each trial a sample stimulus was presented, e.g. A1, followed by two comparison stimuli, e.g. B1 (S+) and B2 (S-). On each trial the sample and comparison (S+) stimuli appeared at full intensity black on white. The incorrect comparison (S-) initially appeared as pale grey on white. As correct responding to S+ was established, the incorrect comparison slowly darkened in colour until it too was at full intensity black on white. There were six different levels of intensity. Three consecutive correct responses were required at each level before the intensity was increased. An incorrect response at any level resulted in the intensity of the S- stimulus immediately being reduced by one level until performance was stable once more. Once both the correct and incorrect comparison stimuli were at full intensity, six consecutive correct responses were required before the subject was said to have reached criterion on that discrimination.

In phase 1, the subjects were taught the relations A1-B1 and A2-B2 using the errorless learning procedure. Performance was then tested on the A-B mix where both discriminations appeared together within a block of trials, with both the sample and both comparison stimuli appearing at full intensity. The subjects received ten trials on the A-B mix with reinforcement, followed by ten trials without reinforcement. The relations B1-C1 and B2-C2 and the B-C mix were then trained and tested in the same way.

Following this, performance on the A-B and B-C relations was reviewed with subjects receiving first six trials with reinforcement and then six trials without reinforcement on each set of relations. In the session testing for equivalence,

maintenance of these relations was reviewed with subjects receiving four trials without reinforcement followed by four trials with reinforcement on each set of relations. This was followed by eight trials on the C-A equivalence relations with no reinforcement available. As with the normally able adults, the aim of this procedure was to establish stable equivalence relations while ensuring that the baseline relations were maintained in the absence of reinforcement.

In phase 2 the relations A1-C2 and A2-C1 were trained using the errorless learning procedure and performance on the A-C mix was tested in the same way as in phase 1. Once again training and testing on the A-C relations and the A-C mix was matched to that required to establish stable performance on the A-B and B-C relations in phase 1.

Performance was then reviewed on all A-B, B-C and A-C relations, with subjects receiving six trials with reinforcement followed by six trials without reinforcement on each set of relations. In the session retesting performance on the C-A equivalence relations, maintenance of each set of relations was tested once more. Subjects received four trials without reinforcement followed by four trials with reinforcement on each set of relations, followed by eight trials without reinforcement on the C-A equivalence relations. This tested for any disruption of the trained baseline conditional discriminations while still ensuring that any disruption to performance on the C-A equivalence relation was not due to extinction in the absence of reinforcement.

Results

1. Normally able adults

The performances shown by the normally able adults are all very similar. All four subjects rapidly reached criterion on the A-B and B-C relations in phase 1. All the subjects maintained high accuracy on these relations in the absence of reinforcement and subsequently demonstrated 100% correct performance on the C-A equivalence relations at the end of phase 1. (See Fig. 10:1; Fig 10:3; Fig. 10:5; Fig. 10:7)

All four subjects made at least one error on the first presentation of the inconsistent A-C relations in phase 2. (See Fig. 10:2; Fig. 10:4; Fig. 10:6; Fig. 10:8). These errors were in keeping with the relations established in phase 1. All subjects then rapidly learned the inconsistent A-C relations and maintained this performance in

the absence of reinforcement. In the session following A-C training which reviewed the A-B, B-C and A-C relations (reinforced trials followed by unreinforced trials) performance was generally highly accurate although two subjects (subjects 3 and 4) made errors on presentations of the B-C or A-C relations (see Fig 10:6; Fig 10:8). In both cases these errors occurred on reinforced trials and both subjects subsequently maintained performance on unreinforced trials. In the unreinforced maintenance check in the session testing for equivalence, three of the subjects showed at least some disruption on the trained relations. Subjects 1 and 4 showed disruption on the A-B, and A-B and B-C relations respectively (see Fig. 10:2; Fig. 10:8). Subject 2's performance on the B-C relations was at chance level while performance on the A-C relations had reversed from that trained so that the subject matched A1-C1 and A2-C2 (see Fig. 10:4). These reversed relations were in accordance with A-C transitivity relations assumed to have been derived in phase 1. All four subjects scored 0% correct on the C-A equivalence retest. As correct performance was taken as C1-A1, C2-A2 as demonstrated at the end of phase 1, this means that all subjects matched C1-A2 and C2-A1. This performance on the C-A equivalence retest is the symmetrical counterpart of the A-C relations trained at the beginning of phase 2.

2. Adults with learning disabilities

Phase 1

All four subjects learned the A-B and B-C relations easily using the errorless learning procedure. Three subjects immediately showed accurate performance on the A-B and B-C mixes (See Fig. 10:11; Fig. 10:13; Fig. 10:15). Subject 5 initially showed very unstable performance on both reinforced and unreinforced presentations of the mixes (see Fig. 10:9). This may have been due to the shift from presenting the relations one at a time (e.g. A1-B1) in the errorless learning procedure, to both relations together (e.g. A1-B1, A2-B2) in the mixes as subsequent performance on these relations was good.

In the session reviewing the A-B and B-C relations performance by all four subjects was highly accurate. In the maintenance check preceding equivalence testing all four subjects showed accurate responding on the A-B and B-C relations. When tested for the emergence of C-A equivalence three subjects immediately showed 100% correct equivalence performance (See Fig. 10:9; Fig. 10:13; Fig 10:15). Subject 6 however, scored 0% on this test, indicating that she had derived the relations C1-A2, C2-A1 and responded consistently on these relations (see Fig. 10:11).

Phase 2

Subject 5 easily learned the A-C relations using the errorless learning procedure. On the first presentation of the A-C mix performance was somewhat unstable (70%) but on the second presentation Subject 5 showed 100% correct responding. In the session reviewing the A-B, B-C and A-C relations subject 5 showed chance level performance on the reinforced trials of the A-B relations and 0% correct performance on the unreinforced relations. His performance on the B-C and A-C relations was generally stable. In the maintenance check preceding equivalence testing performance on the A-B relations was at chance level while performance on the B-C and A-C relations was generally highly accurate. On the C-A retest of equivalence Subject 5 scored 0% correct. As correct performance was defined as the c1-A1, C2-A2 equivalence relations shown in phase 1, Subject 5 matched C1-A2, C2-A1, the symmetrical counterparts of the A-C relations trained in phase 2 (see Fig. 10:10).

Subject 7 also easily learned the A-C relations. Performance on the first presentation of the A-C mix was generally good (80%) and 100% correct on the second presentation. In the session reviewing performance on the a-B, B-C and A-C relations performance was 100% correct on all relations. In the maintenance check preceding equivalence testing subject 7 showed some disruption of performance. performance on the A-B and A-C relations was 100% correct. On the unreinforced B-C relations subject 7 scored 0% correct, indicating that he had in fact reversed the relations taught, selecting B1-C2, B2-C1. On the reinforced trials, performance was 50%, chance level. On the C-A equivalence retest, subject 7 scored 0% correct, indicating that he was selecting C1-A2, C2-A1, the counterpart of the A-C relations trained in phase 2 (see Fig. 10:14).

Subject 8 easily learned the A-C relations. Initially performance on the a-C mix was unstable (60%) but improved on the second presentation (100%). In the session reviewing the A-B, B-C and A-C relations performance was quite unstable. performance on the A-B mix was 100% correct. Initially performance on the B-C relations was 50% correct (chance level) but this improved to 100% correct. performance on the A-C relations was very unstable. Subject 8 initially scored 17% correct on the A-C relations, indicating that subject 8 tended to reverse the trained A-C relations and select A1-C1, A2-C2. On the maintenance check preceding equivalence performance was generally highly accurate with the only errors occurring on the A-C relations (75% correct). On the C-A equivalence retest

subject 8 scored 88% correct, indicating that he almost always responded in accordance with the C-A relations derived in phase 1 (see Fig 10:16).

Subject 6 had failed to show appropriate C-A equivalence relations in phase 1, instead matching C1-A2, C2-A1. She easily learned the A-C relations in phase 2 and displayed generally accurate performance on the A-C mix. In the session reviewing the A-B, B-C and A-C relations performance was generally highly accurate. Subject 6 showed some disruption of performance on the maintenance check preceding the equivalence retest. performance on the A-B mix was 100% correct but on the first, unreinforced presentation of the B-C relations performance was 0% correct, indicating that Subject 6 had reversed the B-C relations, selecting B1-C2, B2-C1. However, on the reinforced B-C trials, Subject 6 reversed *this* performance, scoring 100% correct. Performance on the unreinforced A-C relations was 50% correct but this also improved to 100% correct during the reinforced trials. On the C-A equivalence retest subject 6's performance was 0% correct, matching C2-A1, C1-A2. This was consistent with the inappropriate equivalence relations she had displayed in phase 1 (see Fig. 10:12).

As subject 6 failed to demonstrate appropriate equivalence relations in phase 1 her results will not be included in the discussion of performance in this experiment. However, they indicated that she showed some disruption of performance while trying to integrate the three sets of relations.

Discussion

In this replication of Experiment 4 all the subjects, both normally able and with learning disabilities, were able to respond consistently on the final C-A equivalence retest. This is in contrast to the original experiment where the adult with learning disabilities showed random performance on the equivalence retest.

It was possible that in the original experiment the normally able adults were able to maintain consistent responding by treating the relations trained in phase 1 and the inconsistent relations trained in phase 2 as essentially separate tasks. By not attempting to integrate the two sets of relations it would easily be possible to maintain consistent responding. The procedure in this experiment eliminated this possibility by presenting a review of all trained relations (both phase 1 and phase 2) before retesting equivalence. Even with this procedure the normally able adults were able to maintain consistent responding on the C-A equivalence retest. This

suggests that normally able adults are able to deal with conflicting items of relational information, possibly by developing overt strategies to monitor their responding and maintain performance.

This does not mean that the normally able adults were in any sense able to “solve” the problem of the conflicting information. The relations trained in phases 1 and 2 remain inconsistent, and there is some evidence that the normally able adults experienced difficulty in trying to reconcile these inconsistencies. Three of the normally able adults showed at least some disruption to performance on the maintenance check preceding equivalence retesting, yet none of the subjects showed anything other than 100% correct responding prior to equivalence testing in phase 1.

In this replication, the adults with learning disabilities were also able to maintain consistent responding on the final equivalence test. As the sequence of training and testing was the same as that used with the normally able adults it seems reasonable to conclude that these subjects were also able to develop some overt strategy to monitor their responding and maintain consistent performance. The adults with learning disabilities showed slightly more disruption to performance on the maintenance check prior to retesting equivalence in phase 2. This would suggest that while they were able to deal with the inconsistencies between the relations, they may have found it harder initially to reconcile the inconsistencies.

Only one subject responded according to the equivalence relations derived in phase 1 when retested at the end of phase 2. All the other subjects, both normally able and with learning disabilities matched C1-A2, C2-A1.

It is likely that the ability to cope with inconsistent relations is part of environmental control of equivalence relations. In his 1986 paper Sidman described contextual control of equivalence relations in terms of five-term contingencies. The relations trained and tested in this experiment are all conditional discriminations - four-term contingencies. In this 1986 paper Sidman had suggested that the four-term contingency was the level at which stimulus equivalence emerged. Sidman's 1994 book revised this suggestion, proposing instead that equivalence emerged at the level of the three term contingency, or simple discrimination. (See Chapter 5 for fuller discussion). This reassessment describes the most basic level at which stimulus equivalence may *emerge* and considers the possibility of including responses and reinforcers within equivalence

relations. However environmental control of equivalence relation emerges at the level of the five-term contingency and can be described in that context.

Possibly the most important feature of these results is that so many of the subjects were able to produce consistent responses on the final C-A equivalence retest. In many ways the random responses shown by JD in the original experiment were the most reasonable. The A-B and B-C relations trained in phase 1 and the A-C relations trained in phase 2 are inconsistent and cannot be reconciled. In this case the random responses shown by JD make sense as there is no "correct" solution. one way to deal with this problem would be to treat the relations trained in each phase as separate tasks. However, the review session in phase 2 was designed to prevent this and require the subjects to treat the three sets of relations as connected. In that case it seems likely that the subjects performances can be accounted for in terms of contextual control of the equivalence relations, or five-term contingencies.

Sidman (1986) described how the four-term conditional discrimination established in phase 1 could be placed under control by allowing an additional element of the experimental environment to vary. This is described in terms of the balanced four-term contingency depicted in Fig 5:4 (chapter 5). Fig. 10:17 shows this contingency placed under the contextual control of two further elements, S5 (tone 1) and S6 (tone 2). If tone 1 is sounding then the original contingencies hold true and the subject can get a reward by pressing the square in the presence of the green hue, and by pressing the circle in the presence of the red hue. However, S6 (tone 2) causes these contingencies to reverse, so that a reward is obtained for pressing the circle in the presence of the green hue and the square in the presence of the red hue. The tones assume contextual control over the original conditional discrimination.

It is likely that a similar process operated in this experiment. Phase 1 training established a balanced conditional discrimination , through stimulus equivalence, so that on the final equivalence test the subjects selected A1 in the presence of C1, and A2 in the presence of C2. The phase 2 training introduced inconsistent conditional discriminations which the subjects were required to try and reconcile with the original discrimination. The "strategies" which allowed consistent responding on the final C-A equivalence retest may be due to the development of contextual control.

All three sets of trained relations were reviewed in phase 2 and there is some evidence that the subjects tried to integrate them. For instance, subjects 2, 5, 7 and

8 all reversed one of the trained relations at one stage within phase 2. Reversing one set of relations would allow the subject to integrate the three sets of trained relations. For instance, subject 7 reversed the B-C relations on the unreinforced B-C maintenance check prior to retesting C-A equivalence, thus selecting B1-C2, B2-C1. This would have produced the equivalence classes A1B1C2, A2B2C1 and in this way reconciled the relations trained. However, these unreinforced trials were followed by reinforced trials where the subject was reinforced for selecting the relations B1-C1, B2-C2, thus preventing this way of reconciling the inconsistent relations. What may have happened is that contextual control based on the phase 1 or phase 2 relations may have developed. Thus, if the C-A equivalence test was preceded by trials of the relations trained in phase 1 then the subjects might have matched C1-A1, C2-A2 on the equivalence test. However, if the equivalence test was preceded by trials of the relations trained in phase 2 then the subjects might have matched C1-A2, C2-A1.

In this experiment the C-A equivalence retest was always preceded by reinforced trials of the A-C relations trained in phase 2. All but one of the subjects tested matched C1-A2, C2-A1, suggesting that these A-C trials had established contextual control determined by the relations trained in phase 2. However, one subject (subject 8) almost always matched c1-A1, C2-A2 when C-A equivalence was retested. It is still possible though, that this subject's responding was also under contextual control. The only time the C stimuli had functioned as samples and the A stimuli as comparisons was in the context of the equivalence test at the end of phase 1. Thus, the stimuli appearing in this arrangement may have been sufficient to establish contextual control according to the contingencies established in phase 1, and so the subject matched C1-A1, C2-A2.

The five term contingency described by Sidman (see Fig. 10:17) clearly defines the stimuli established as contextual control (S5 - tone 1; S6 - tone 2). However, no contextual cue was clearly defined for this experiment, and so the context would have depended on the cue which appeared most salient for each subject. For most subjects this appeared to be the phase 2 A-C relations immediately preceding testing but for subject 8 it appears to have been the C-A equivalence test itself.

There are several ways in which it might be possible to test this proposition. For instance, it would be possible to vary the order in which the training relations are presented prior to the retesting of C-A equivalence. Thus, if trials on the phase 1 relations preceded testing for some subjects, and phase 2 relations preceded testing

for others, this might establish different contextual cues for performance on the C-A equivalence retest. Alternatively, it would be possible to provide explicit contextual cues, such as the tones suggested by Sidman, while training the phase 1 and phase 2 relations. The presentation of one of these cues while retesting C-A equivalence might be sufficient to determine performance regardless of the order of training and testing.

The process of establishing contextual control, as suggested here, seems likely to be an important part of human behaviour. Sidman (1994) pointed out that "Without second-order conditional control we would be seriously handicapped in adapting to the complex environment in which we live" (p.409). Conditional discriminations in the natural environment do not often fall into neat, consistent equivalence relations, and a relation which holds true in one context may not hold true in another.

Thus, while the relations presented cannot be resolved neatly, and random responding may make *logical* sense, the establishment of contextual control can be seen as an adaptive response. "Given that reinforcement contingencies do create five-term units, we have to conclude that the incompatibility is usually resolved in favor of differential control" (Sidman, 1994, p.409).

Given this explanation of the results obtained, the failure of JD to show consistent responding in the original experiment also makes sense. The training and test procedure does not appear to have been sufficient to establish contextual control over her responding. Given that there is evidence that she tried to integrate the inconsistent relations, this would explain her random performance on the final test. Without a "strategy" such as contextual control to govern performance it would be extremely difficult to produce a consistent response. Therefore, while JD appears able to derive equivalence relations appropriately, she may find difficulty in manipulating these relations according to context.

Similarly, the failure of the normally able children to respond consistently on the equivalence retest during the original experiment may be due to a lack of experience of contextual control. The children were all able to demonstrate appropriate equivalence class formation but, to respond consistently on the final retest would have required them to establish contextual control by themselves. They may simply have lacked sufficient experience of equivalence relations switching according to context to establish this control themselves.

Table 10:1 Training and Test Stimuli

Stimulus Class	A	B	C
----------------	---	---	---

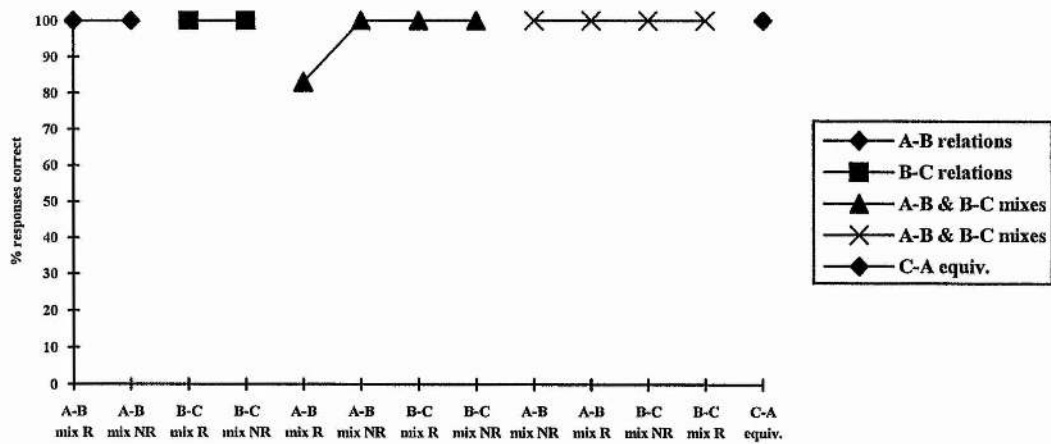
Normally able adults and young children

1	Ξ	ξ	xi
2	Φ	ϕ	phi

Adult with learning disabilities

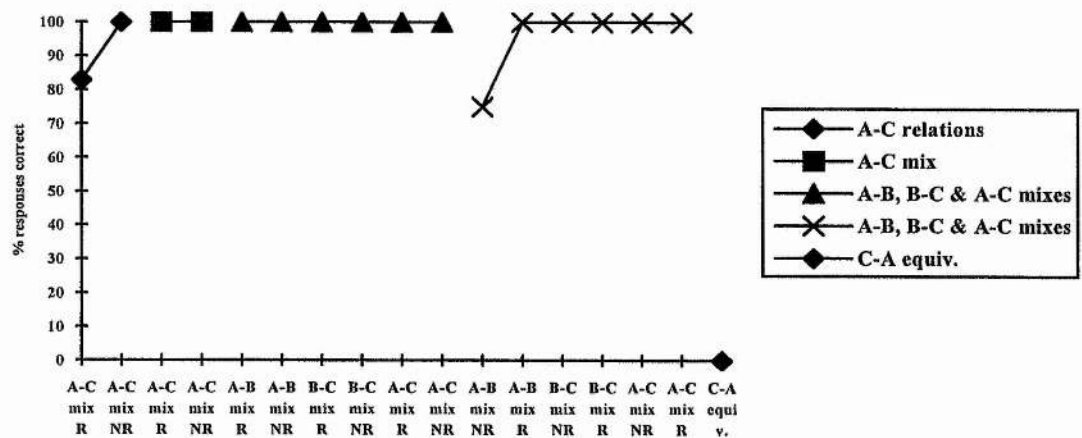
1	Γ	γ	gamma
2	Δ	δ	delta

Fig. 10:1 Subject 1- Phase 1 results



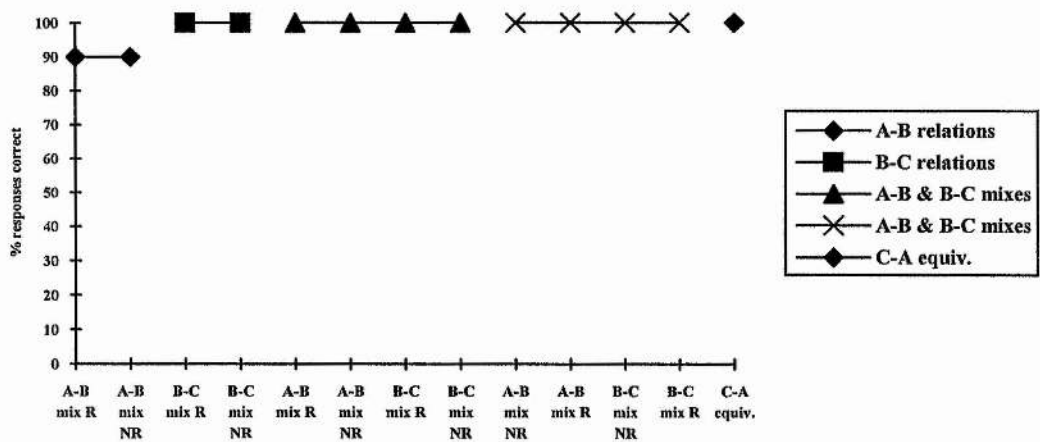
R - denotes reinforced trials
 NR - denotes no reinforcement

Fig. 10:2 Subject 1 - Phase 2 results



R - denotes reinforced trials
 NR - denotes no reinforcement

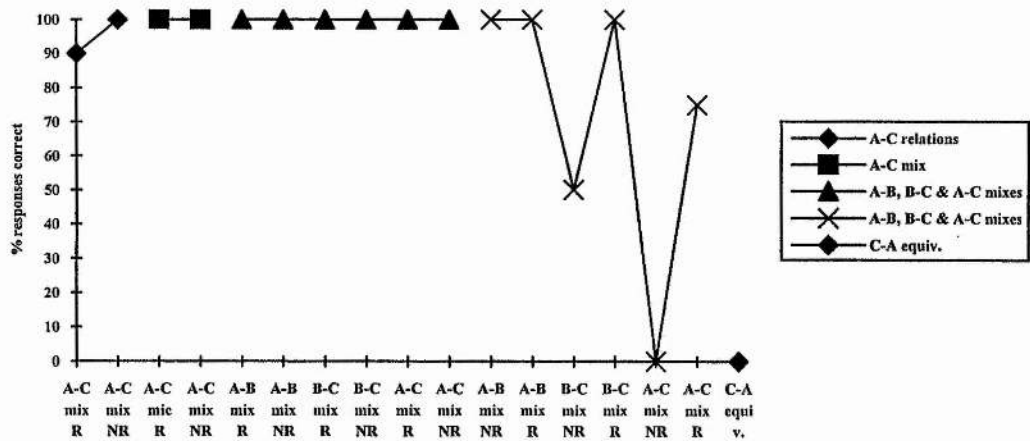
Fig. 10:3 Subject 2 - Phase 1 results



R - denotes reinforced trials

NR - denotes no reinforcement

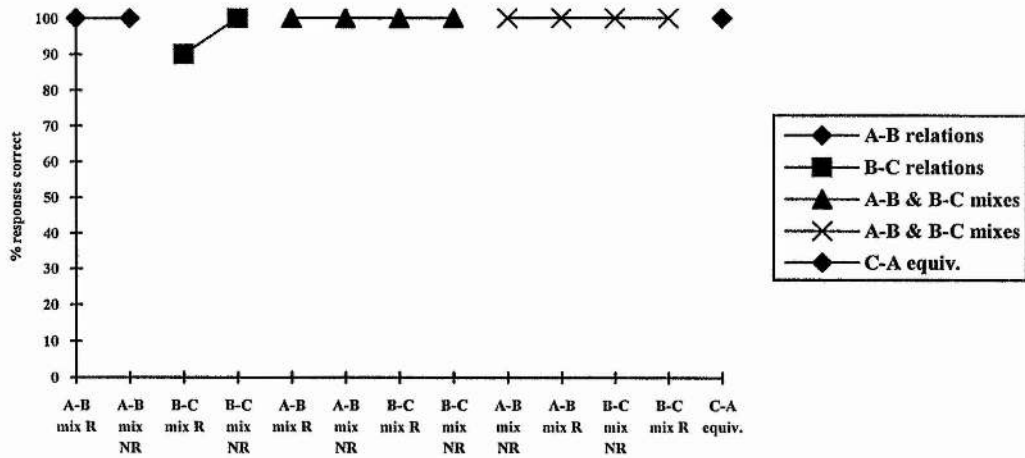
Fig. 10:4 Subject 2 Phase 2 results



R - denotes reinforced trials

NR - denotes no reinforcement

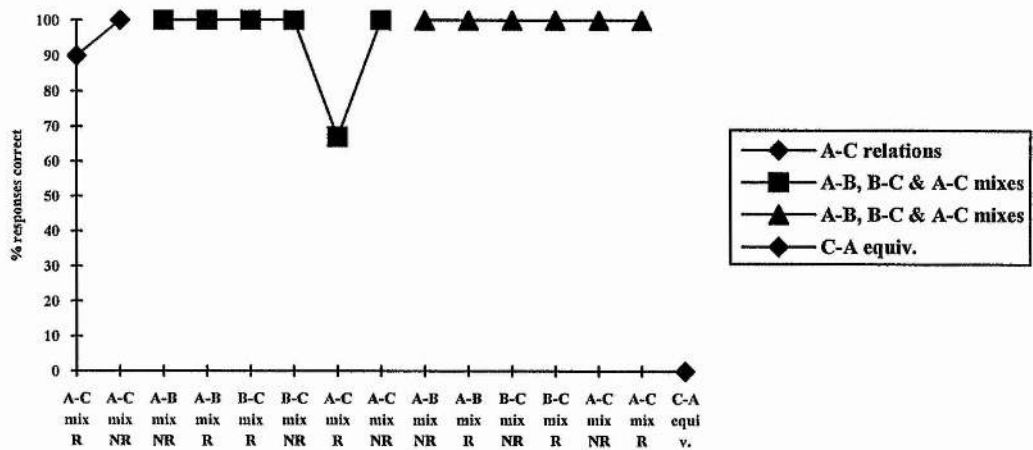
Fig 10: 5 Subject 3 - Phase 1 results



R - denotes reinforced trials

NR - denotes no reinforcement

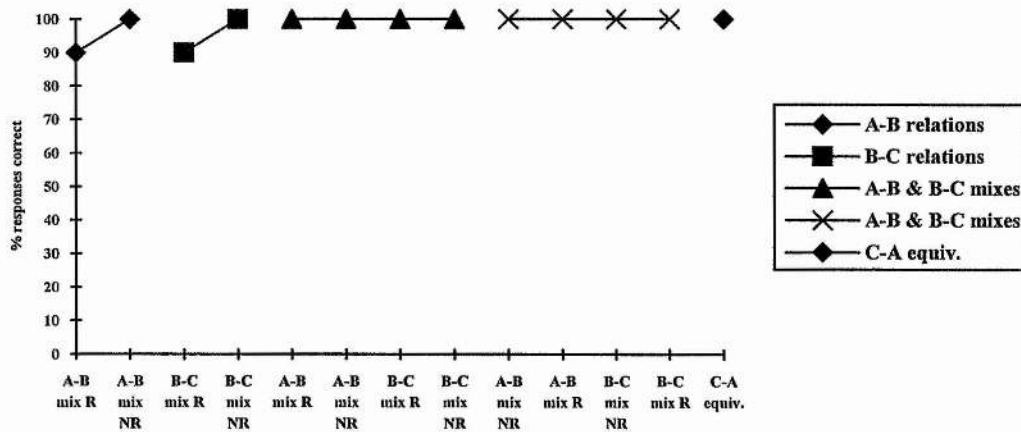
Fig. 10:6 Subject 3 - Phase 2 results



R - denotes reinforced trials

NR - denotes no reinforcement

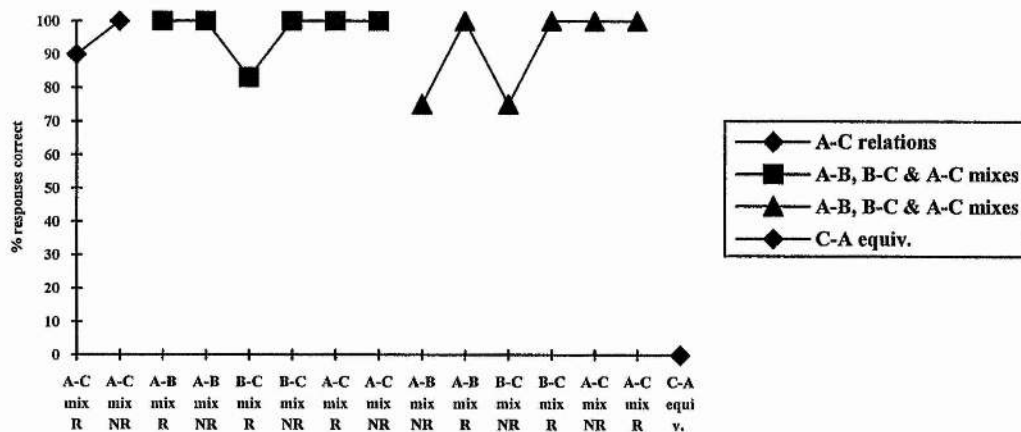
Fig 10:7 Subject 4 - Phase 1 results



R - denotes reinforced trials

NR - denotes no reinforcement

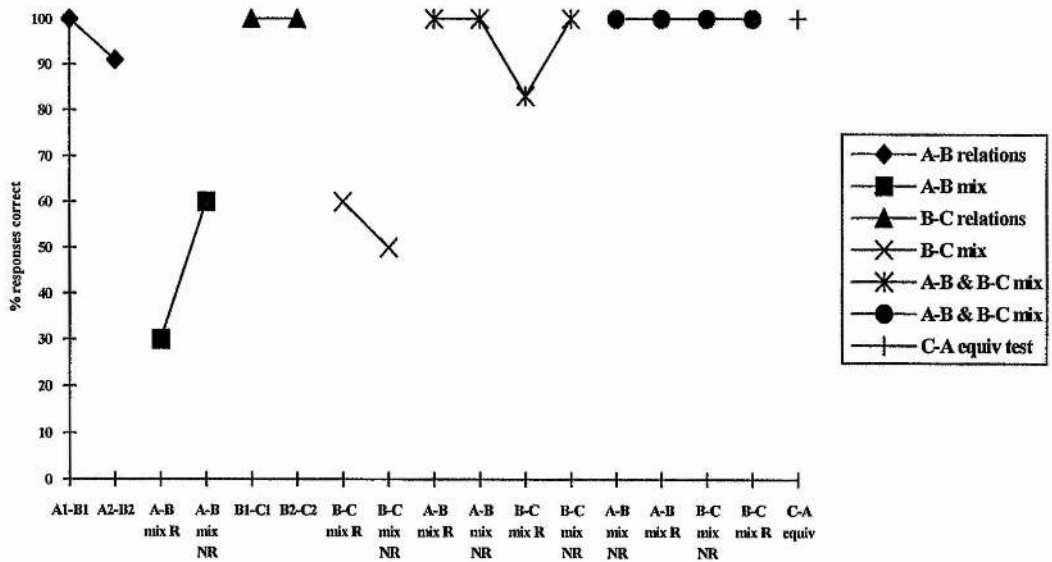
Fig 10:8 Subject 4 - Phase 2 results



R - denotes reinforced trials

NR - denotes no reinforcement

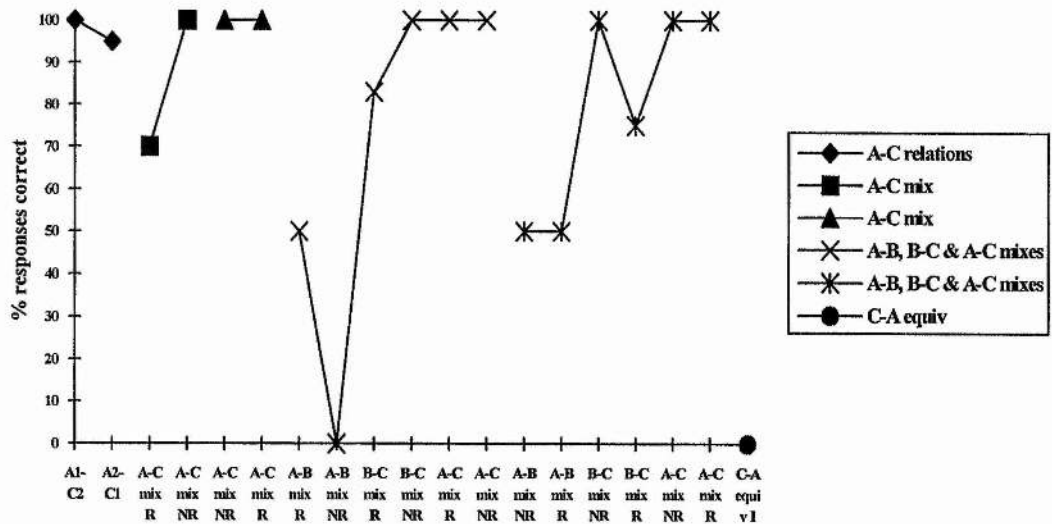
Fig. 10:9 Subject 5 - Phase 1 results



R - denotes reinforced trials

NR - denotes no reinforcement

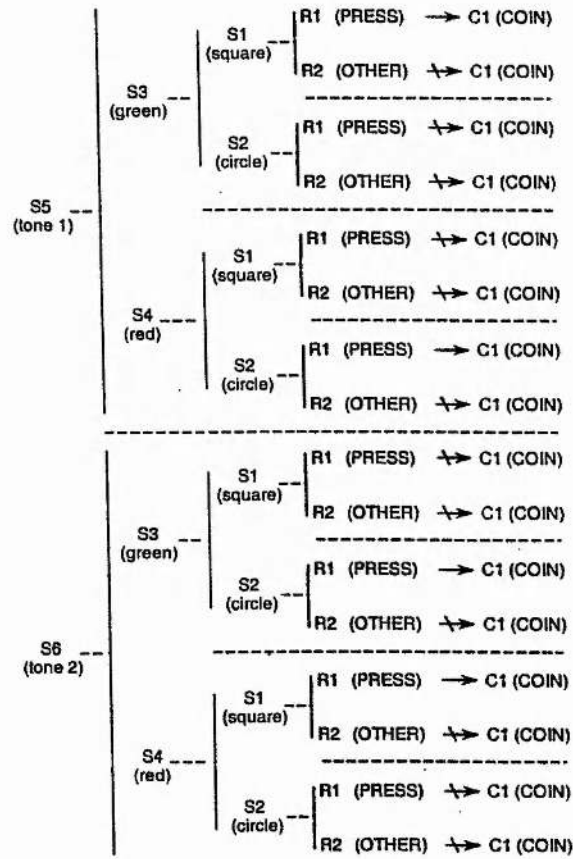
Fig. 10:10 Subject 5 - Phase 2 results



R - denotes reinforced trials

NR - denotes no reinforcement

Fig. 10:17 The five-term contingency (Second-Order Conditional Discrimination)
 S = Stimulus; R = Response; C = Consequence
 From Sidman (1986)



CHAPTER 11

Discussion

The five experiments presented in this thesis describe investigations of specific aspects of stimulus equivalence and transitive inference with several groups of subjects. The question then, is what do the results from these experiments show about the establishment of emergent relations in general, and specifically, their establishment with subjects with learning disabilities ?

Emergent Relations Require the Establishment of
More Than the Baseline Relations

One feature which was particularly notable from experiments 1 and 2 is that, for the emergence of both stimulus equivalence and transitive inference, it is not sufficient merely to teach the baseline relations to criterion. For instance, in experiment 1, four subjects received training on the stimulus relations A1-B1, A2-B2, A1-C1 and A2-C2, and were tested for the emergence of B-C and C-B equivalence relations. While three out of four subjects reached criterion on these baseline A-B and A-C relations, only one subject, JD, immediately demonstrated equivalence relations. Subject NL received extensive training on the baseline conditional discriminations but failed to display any of the components of equivalence. Similarly, subject AS (case study: Appendix I) reached criterion on the A-B and B-C relations, but failed to demonstrate equivalence on two occasions. It was also clear that AS had not derived any of the properties of equivalence, as even when the A-B baseline relations were at criterion, performance on the B-A test of symmetry was at chance level.

In experiment 2, subject MM failed to demonstrate transitive inference responding even though he had reached criterion on the baseline training pairs. However, it does seem likely that this failure to demonstrate transitive inference may be due to MM's failure to maintain performance on these baseline relations in the absence of reinforcement as he did demonstrate transitive inference responding in experiment 4 when performance on the baseline relations was maintained. Nevertheless, it would appear that merely accurate performance on the baseline relations is not sufficient for the emergence of untaught behaviour. These new emergent relations have never been directly involved in any reinforcement contingency when demonstrated in an experimental setting, yet with some subjects these relations are demonstrated as positively as the directly reinforced baseline relations.

The question of precisely what it is that causes these relations to emerge has been strongly debated within the context of equivalence research. In this thesis, performance by adults with learning disabilities on an equivalence task was compared to performance on an inference task. The aim was to see if the comparison with another example of emergent relations would give some indication of the processes underlying stimulus equivalence. It was hypothesised that if a subject demonstrated stimulus equivalence, that subject should also be able to demonstrate transitive inference, as there seemed to be some evidence that equivalence relations were more difficult to derive than inference relations. This was based on the greater difficulty of demonstrating equivalence, compared to inference, with non-humans. This hypothesis was supported by the fact that JD who immediately demonstrated equivalence relations in experiment 1, also clearly demonstrated transitive inference in experiment 2.

It was also hypothesised that subjects who failed to demonstrate transitive inference would consequently also fail to demonstrate stimulus equivalence, and that subjects who did demonstrate transitive inference would not necessarily also be able to demonstrate stimulus equivalence. Again, this was because it appears easier to demonstrate transitive inference than stimulus equivalence with non-human subjects, suggesting that transitive inference may be a more basic skill. Unfortunately it is not clear that these hypotheses were supported as neither subjects GT or TD who demonstrated stimulus equivalence in experiment 1 were available for transitive inference testing, and subject MM who participated in the study of transitive inference performance in experiment 2 was not tested on stimulus equivalence.

Possible Concrete vs. Abstract Relations Distinction

Experiments 3, 4 and 5 took a different approach to examining the processes operating in inference and equivalence. Each experiment established stable performance on one of the paradigms and then tested the effects of a disruption of the baseline relations, and how this affected the performance of three groups of subjects, normally able adults, adults with learning disabilities, and normally developing young children. In both experiments 3 and 4 the normally able adults were able to maintain consistent responding after the baseline relations were disrupted, while the adults with learning disabilities showed considerable disruption of performance following the introduction of the inconsistent relation.

On the transitive inference task, the normally developing young children were generally able to maintain consistent responding, while the young children given the disrupted equivalence task did not. In experiment 5, both the normally able adults and the adults with learning disabilities were able to maintain consistent responding. In all these experiments, when the inconsistent relation was introduced, the original baseline relations no longer made logical sense. This may explain why the normally able adults were able to maintain responding while some of the adults with learning disabilities were not. In a similar manipulation of a transitive inference task, Siemann and Delius (1994) had found that normally able adults were able to maintain responding, but that they used a non-transitive strategy when they did so. It is possible that the major difference between the transitive inference and stimulus equivalence paradigms is that the relations in transitive inference are based on some concrete dimension (e.g. "greater than" vs. "lesser than") while the relations in stimulus equivalence are completely arbitrary ("goes with"). [A.C. Catania (personal communication, December 1993)]

Another way of looking at this is that in both these paradigms (transitive inference - experiment 3 and stimulus equivalence - experiments 4 and 5) the introduction of an inconsistent relation in phase 2 added another dimension to the training already given and the subjects with learning disabilities may have had more difficulty coping with this shift in complexity. In effect it turned the baseline relations from three-term to four-term contingencies in transitive inference, and from four-term to five-term contingencies in stimulus equivalence. Both MM and JD, who showed disruption to their established performance, might have been able to cope with these contingencies if they had been present from the start of training but may have had difficulty coping with this development within the experiment.

In experiment 3, the phase 1 training established a consistent transitive inference series, but introducing the inconsistent relation in phase 2 destroyed this, effectively removing the concrete dimension and making the relations arbitrary. In effect turning the pair comparisons from three-term simple discriminations to four-term conditional discriminations. The normally able adults are likely to have much more experience with arbitrary/abstract relations, and thus be able to bring in non-transitive strategies to deal with the alteration to relations. However, the subjects with learning disabilities will probably have much less experience with relations of this type, and therefore may not have strategies available to deal with these relations, or at least may find it difficult to switch strategies during testing.

In experiments 4 and 5, while the relations between the stimuli in phase 1 were arbitrary equivalence relations, they were also logical and consistent. The training in phase 2 then established relations inconsistent with the previous training and the relations, as presented, stopped making logical sense. The normally able adults in both experiments 4 and 5 and the adults with learning disabilities in experiment 5 were able to deal with this shift in complexity. The training in phase 2 of the experiments seemed to cause the relations already established to come under contextual control - the two contexts being either phase 1 or phase 2 training. Subject JD in experiment 4 did not appear to develop this contextual control of responding. Possibly her level of learning disability was more severe than that of the adults with learning disabilities in experiment 5. She may have had less experience of conditional discriminations coming under contextual control and thus not been able to derive this for herself. or possibly the training procedure in experiment 5, with its frequent reviews of the relations presented in phases 1 and 2, was more likely to produce contextual control of the conditional discriminations. JD, the subject with learning disabilities in experiment 4, was clearly able to deal with arbitrary relations, as she had demonstrated stimulus equivalence on two occasions. However, for whatever reason, it seems that it was more important for her than for the other adult subjects that the trained relations preserved their logical structure.

The normally developing children in experiment 3 were able to maintain responding on the transitive inference test pairs even after the introduction of the inconsistent relation. The children in this experiment had a mean age of 5 years and 4 months, had all experienced at least one year of formal schooling, and were able to read and write at levels consistent with their chronological ages. Artman and Cahan (1993) had suggested that schooling is one of the major influences on the development of transitive inference skills, and Bryant and Trabasso (1971) had demonstrated reliable transitive inference performance with children at least a year younger than those tested in this experiment. It was likely then that these children would have well developed transitive inference skills and would have experience of quite abstract tasks. This would fit with their ability to maintain responding on the inconsistent transitive inference series.

The normally developing young children in experiment 4 were considerably younger than those in experiment 3, with a difference of nearly 15 months in the mean ages between the two groups. None of these children had received formal schooling at the time of testing. The subjects were clearly able to cope with

arbitrary relations between the stimuli as they demonstrated equivalence relations in phase 1 of the study. However none of these subjects were able to maintain consistent responding after the introduction of the inconsistent relation, although the performance of subject 2 gave some indication that he might be able to respond consistently. As with JD, the subject with learning disabilities in this experiment, this may be because the subjects had little experience with very abstract relations, and thus either did not have alternative strategies available to deal with this task, or were unable to switch strategies during the experiment.

Thus, transitive inference and stimulus equivalence may just describe concrete and arbitrary forms of general emergent behaviour, the relations in transitive inference varying on some sort of scale, while the relations in stimulus equivalence describe an arbitrary "goes with" relation. It may be that experience with either of these paradigms would influence the development of the performance on the other paradigm if both are examples of general emergent relations. Therefore, if a subject appears to have difficulty in deriving these emergent relations, then the transitive inference task which appears to be based on a concrete dimension may provide a useful way of giving experience with these relations, and thus facilitate the emergence of more abstract relations of equivalence. This does not mean that transitive inference relations are a necessary pre-requisite for emergence relations, just that they may provide an accessible way of giving experience with relations from which further relations can be derived.

Applications of This Research

If a subject does have difficulty in demonstrating emergent relations, of whatever type, it is surely valuable to try and find some way of facilitating the emergence of these relations. In the final chapter of his book on equivalence relations, Murray Sidman (1994) commented that one of the disappointments of equivalence research is that it has appeared to have little effect on general education methods.

"Techniques that succeed in teaching so much even to supposedly unteachable people have become standard in the laboratory but have not been adapted by those whose major responsibility is to carry out instruction outside the laboratory" (p.532).

It is disappointing that the training procedures for stimulus equivalence have not been more widely adopted, as numerous experiments have demonstrated the potential power of this teaching paradigm. Even a standard equivalence task which directly trains the relations A-B and B-C results in the emergence of seven further

untaught relations; three of reflexivity, A-A, B-B, C-C, two of symmetry, B-A, C-B, one of transitivity, A-C, and one of equivalence C-A. Many of the early equivalence experiments described the emergence of these relations with subjects with quite severe learning disabilities. These experiments described the emergence of quite complicated relations, of which the subject did not appear to be capable prior to training, yet following equivalence training could demonstrate reliably. If relations of the type described by stimulus equivalence are shown to be crucial to natural learning processes, then it is particularly important that ways are found to establish these relations effectively with subjects who do not appear to demonstrate them spontaneously. As Sidman (1994) pointed out, this paradigm could be even more powerful in a teaching situation than in a laboratory situation. In a teaching situation it is not necessary to demonstrate that the relations would be derived solely as a result of the baseline relations trained. The baseline relations would not require such a rigorous learning criterion, the training and test tasks could be mixed throughout teaching, and reinforcement could be provided on both training and test trials. In a teaching situation the focus is on reliably establishing equivalence relations rather than examining the precise means by which they are derived.

In this context, the question of whether equivalence is a fundamental stimulus function (cf. Sidman 1990, 1994), a product of naming (cf. Dugdale and Lowe 1990, Horne and Lowe, in press), or an example of arbitrarily applicable relational responding (cf. Hayes 1991), becomes of secondary importance, although it would probably be easier to develop appropriate training procedures if the basis of stimulus equivalence was clear. However, this is not a necessity in order for stimulus equivalence to be a useful training procedure. In experiment 1, two subjects, TD and GT did not initially demonstrate equivalence relations, despite reaching criterion on the baseline training relations. After being taught to name the stimuli during conditional discrimination training both subjects demonstrated appropriate equivalence responding. It may be that equivalence emerges as a result of naming as Dugdale and Lowe (1990) and Horne and Lowe (in press) have suggested, and that this explains why teaching naming resulted in equivalence class formation with these two subjects. Or, it may be that naming is not the natural basis of stimulus equivalence in normally able subjects, but does constitute an effective method of establishing equivalence relations with subjects who have difficulty displaying these relations spontaneously. While naming may not be the basis of stimulus equivalence it may provide a useful tool for establishing equivalence relations.

In people with physical disabilities, many of the physical aids used do not precisely mimic the natural physical processes they support. A wheelchair after all provides a very different way of getting around compared to walking, but the aim is to provide a functional alternative when the natural physical process is not possible. Similarly, with subjects with learning disabilities, it may be more appropriate to focus on the most effective ways of helping the subject derive emergent relations, rather than trying to precisely reproduce these relations as they would normally be derived.

Clearly it is helpful if the procedures used to establish equivalence relations are close to those found naturally, as this increases the likelihood of bringing into action as many as possible of the subject's natural learning abilities. However, as equivalence has been shown to be such a potentially powerful teaching tool, the focus should be on establishing these relations in the most effective way possible.

General Conclusions

It is in this context that a theoretical analysis of equivalence relations and an applied analysis of these relations can provide insight into equivalence relations and emergent behaviour in general. A clearer understanding of the processes operating in equivalence can help in the development of effective ways of establishing those relations. Similarly, description of effective ways of establishing these relations may provide useful information on processes operating in equivalence class formation. These investigations are important for subjects with learning disabilities, such as those taking part the experiments described in this thesis. This includes not only subjects like TD and GT in experiment 1, who demonstrated equivalence relations after a fairly simple manipulation of the training paradigm, but also subjects like NL, who did not reach criterion for testing in either experiment 1 or experiment 2. Given NL's difficulty in learning the baseline relations, it is likely that she in particular would benefit from a really effective equivalence teaching procedure, as she is likely to have considerable difficulty in spontaneously demonstrating equivalence relations. The results from experiments 3 and 4 suggested that even when adults with learning disabilities do demonstrate emergent untaught relations, these performances may be less robust than those displayed by normally able subjects, and so any procedure which helps maintain stable performance on these tasks is likely to be beneficial. In this way the stimulus equivalence paradigm may constitute a useful tool for developing emergent relations with subjects who have difficulty displaying these relations spontaneously.

REFERENCES

- Adams, B.J., Fields, L. & Verhave, T. (1993)
Effects of test order on intersubject variability during equivalence class formation.
The Psychological Record, 1993, 43, 133-152
- Aeschelman S.R. & Higgins A.F. (1982)
Concept learning by retarded children : a comparison of three discrimination learning procedures.
Journal of mental deficiencies research, 1982, 26, 229-238
- Artman, L. & Cahan, S. (1993)
Schooling and the development of transitive inference
Developmental Psychology, 1993, 29(4), 753-759
- Barnes, D. (1994)
Stimulus equivalence and relational frame theory.
The Psychological Record, 1994, 44, 91-124.
- Barnes, D, McCullagh, P.D. & Keenan, M. (1990)
Equivalence class formation in non-hearing impaired children and hearing impaired children
The Analysis of Verbal Behavior, 1990, 8, 19-30
- Braine, M.D.S. (1959)
The ontogeny of certain logical operations: Piaget's formulation examined by non-verbal methods.
Psychological Monographs, 1959, 73(5, Whole No. 475)
- Braine, M.D.S. (1963)
The ontogeny of English phrase structure: The first phrase.
Language, 1963, 39, 1-13
- Braine, M.D.S. (1964)
Development of a grasp of transitivity of length: A reply to J. Smedslund.
Child Development, 1964, 35, 779-810

Breslow, L. (1981)

Reevaluation of the literature on the development of transitive inferences.

Psychological Bulletin, 1981, 89(2),325-351

Brown, R. (1973)

A first language: the early stages.

Cambridge: Harvard University Press, 1973.

Bricker, D.D. (1972)

Imitative sign training as a facilitator word-object association with low functioning children.

American Journal of Mental Deficiency, 1972, 76, 509-516

Bryant, P.E. & Trabasso, T. (1971)

Transitive inferences and memory in young children

Nature, 1971, 232, 456-458

Boysen, S.T. & Berntson, G.G. (1989)

Numerical competence in a chimpanzee (*Pan troglodytes*)

Journal of Comparative and Physiological Psychology, 103, 23-31

Burt, C.J. (1919)

Journal of Experimental Pedagogy, 5, 68-77

Cahan, S. & Cohen, N. (1989)

Age versus schooling effects on intelligence development.

Child Development, 60, 1239-1249

Cahan, S. & Davis, D. (1987)

A "between grade levels" approach to the investigation of the absolute effects of schooling on achievement

American Educational Research Journal, 24, 1-12

Catania, A.C. (1986)

On the differences between verbal and nonverbal behavior.

Analysis of Verbal Behavior, 4, 2-9

Ceci, S.J. (1991)

How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence.

Developmental Psychology, 27, 703-722

Clark, H.H. (1969)

Linguistic processes in deductive reasoning.

Psychological Review, 1969, 76, 387-404

D'Amato, M.R. & Colombo, M. (1988)

Representation of serial order in monkeys (*Cebus apella*)

Journal of Experimental Psychology: Animal Behavior Processes, 14, 131-139

D'Amato, M.R., Salmon, D.P., Loukas, E., & Tomie, A. (1985)

Symmetry and transitivity of conditional relations in monkeys (*Cebus apella*) and pigeons (*Columba livia*)

Journal of the Experimental Analysis of Behavior, 1985, 44(1), 35-47

Davis, D. (1992)

Transitive inference in rats (*Rattus norvegicus*)

Journal of Comparative Psychology, 1992, 106(4), 342-349.

Deese, J. (1965)

The structure of association in language and thought.

Baltimore: The Johns Hopkins University Press, 1965.

Devany, J.M., Hayes, S.C. & Nelson, R.O. (1986)

Equivalence class formation in language-able and language-disabled children

Journal of the Experimental Analysis of Behavior, 1986, 46, 243-257

Dickens, D.W., Bentall, R.P. & Smith, A.B. (1993)

The role of individual stimulus names in the emergence of equivalence relations : the effects of interpolated paired associates training of discordant associations between names.

The Psychological Record, 1993, 43, 713-724.

- Dixon, M. & Spradlin, J. (1976)
Establishing stimulus equivalences among retarded adolescents
Journal of Experimental Child Psychology, 21, 144-164
- Dube, W.V., McIlvane, W.J., Callahan, T.D., & Stoddard, L.T. (1993)
The search for stimulus equivalence in non-verbal organisms.
The Psychological Record, 1993, 43, 761-778.
- Dugdale (1988)
The role of naming in stimulus equivalence: Differences between humans and animals
Unpublished Ph.D. thesis, University of Wales
- Dugdale, N. A., & Lowe, C.F. (1990)
Naming and stimulus equivalence
In D.E. Blackman & H. Lejeune (Eds.)
Behaviour Analysis in theory and practice: contributions and controversies.
Hove, England: Erlbaum.
- Eikeseth, S. & Smith, T. (1992)
The development of functional and equivalence classes in high-functioning autistic children: the role of naming.
Journal of the Experimental Analysis of Behavior, 1992, 58, 123-133
- Fersen, L. von, Wynne, C.D.L., Delius, J.D. & Staddon, J.E.R. (1991)
Transitive Inference Formation in pigeons
Journal of experimental Psychology: Animal behaviour processes, 1991, 17(3), 334-341
- Gillan, D.J. (1981)
Reasoning in the chimpanzee: II. Transitive inference.
Journal of Experimental Psychology: Animal behaviour Processes, 1981, 7(2), 150 - 164.
- Gould, I. (1962)
Perception
In A.J. Bachrach (Ed.),
Experimental foundations of clinical psychology.
New York: Basic Books, 1962.

Hayes, S.C.(1989)

Nonhumans have not yet shown stimulus equivalence

Journal of the Experimental Analysis of Human Behavior, 1989, 51, 385-392

Hayes, S.C. (1991)

A relational control theory of stimulus equivalence

In L.J. Hayes & P.N. Chase (Eds.)

Dialogues on verbal behavior

Reno, NV: Context Press

Hayes, S.C. & Hayes, L.J. (1989)

The verbal action of the listener as a basis for rule governance.

In S.C. Hayes (Ed.)

Rule-governed behavior: Cognition, contingencies, and instructional control.

New York: Plenum

Hayes, S.C. and Hayes, L.J. (1992)

Verbal Relations and the evolution of Behavior Analysis

American Psychologist, 1992, 47,(11), 1383-1395

Higa, J.J. & Staddon, J.E.R. (1993)

"Transitive Inference" in multiple conditional discriminations

Journal of the Experimental Analysis of Human Behavior, 1993, 59, 265- 291

Hird, J. (1989)

Stimulus equivalence in mentally handicapped and normal subjects: The role of language

Unpublished Ph.D thesis.University of Wales, Bangor, U.K.

Hogan, D.E. & Zentall, T.R. (1977)

Backward associations in the pigeon.

American Journal of Psychology, 1977, 90, 3-15

Holmes, P.W. (1979)

Transfer of matching performance in pigeons.

Journal of the Experimental Analysis of Behavior, 1979, 31, 103-114

Horne, P.J. & Lowe, C.F. (in press)

On the origins of naming and other symbolic behavior

Journal of the Experimental Analysis of Behavior

Hutchinson, J. (1986)

Children's sensitivity to the contrastive use of object category terms.

Papers and Reports on Child Language Development, 25, 49-56

Jenkins, J.J. (1963)

Mediated associations: Paradigms and situations.

In C.N. Cofer & B.S. Musgrave (Eds.)

Verbal Behavior and learning: Problems and processes.

New York: McGraw-Hill, 1963.

Jenkins, J.J. (1965)

Mediation theory and grammatical behavior.

In S. Rosenberg (Ed),

Directions in psycholinguistics.

New York: Mcmillan, 1965, 66-96.

Jenkins, J.J. & Palermo, D.S. (1964)

Mediation processes and the acquisition of linguistic structure.

In U. Bellugi, & R. Brown (Eds.)

The acquisition of language.

Monographs of the Society for Research in Child Development, 1964, 29, Serial No. 92, 141-169

Jones L. & Cullen C. (1980)

Errorless learning with the retarded : some further considerations

Journal of practical approaches to developmental handicap, 1980, 4(2), 19-23

Kastak, D. & Schusterman, R.J. (1993)

Immediate transfer of visual identity matching-to-sample in two California sea lions
(Zalophus californianus)

Manuscript submitted for publication

Kendall, S.B. (1983)

Tests for mediated transfer in pigeons

The Psychological Record, 1983, 33, 245-256

Kohler, W. (1927)

Mentality of Apes (in English translation),

London, England; Methuen

In Chance, M.R.A. (1960)

"Kohler's chimpanzees - How did they perform?"

In A.J. Riopelle (Ed.)

Animal Problem Solving (1967)

Middlesex, England; Penguin Books

Kuczaj, S.A., & Donaldson, S.A. (1982)

If the Boy Loves the Girl and the Girl Loves the Dog, Does the Boy Love the Dog?:

The Overgeneralisation of Verbal Transitive Inference Skills

Journal of Psycholinguistic Research, 1982, 11(3), 197-206

Kawashima, T. (1993)

Personal communication to M. Sidman

In Sidman, M. (1994)

Equivalence relations and behavior: A research story

Boston, MA: Authors Cooperative, Inc.

Lazar, R. (1977)

Extending sequence-class membership with matching to sample

Journal of the Experimental Analysis of Behavior, 1977, 27(2), 381-392

Lazar, R.M. & Kotlarchyk, B.J. (1986)

Second-order control of sequence-class equivalences in children

Behavioural Processes, 1986, 13, 205-215

Lipkens, R., Kop, P.F.M., & Matthijs, W. (1988)

A test of symmetry and transitivity in the conditional discrimination performances of pigeons

Journal of the Experimental Analysis of Behavior, 1988, 49(3), 395-409

- Lipkens, R., Hayes, S.C. & Hayes, L.J. (1993)
Longitudinal Study of the Development of Derived Relations in an Infant
Journal of Experimental Child Psychology, 1993, 56(2), 201-239
- Lloyd Morgan, C. (1909)
Introduction to comparative psychology
In A.J. Riopelle (Ed.)
Animal Problem Solving (1967)
Middlesex, England; Penguin Books
- Lovaas, O.I. (1977)
The autistic child: Language development through behavior modification.
New York: Irvington Publishers
- Lowe, C.F. & Beasty, A. (1987)
Language and the emergence of equivalence relations: A developmental study.
Bulletin of the British Psychological Society, 40, A42
- Luiselli J.K. & Donellon S. (1980)
Use of a visual stimulus fading procedure to teach color naming to an autistic child
Journal of Behaviour Therapy & Experimental Psychiatry, 1980, 11, 73-76.
- Lutkus, A. & Trabasso, T. (1974)
Transitive inferences by pre-operational retarded adolescents
American Journal of Mental Deficiency, 1974, 78(5), 559-606
- McGonigle, B.O. & Chalmers, M. (1977)
Are monkeys logical ?
Nature, 1977, 267, 694-696
- McIlvane, W.J. & Stoddard, L.T. (1981)
Acquisition of matching-to-sample performances in severe retardation: Learning by exclusion.
Journal of Mental Deficiency Research, 25, 33-48
- McIntire, K.D., Cleary, J. & Thompson, T. (1987)
Conditional relations by monkeys: reflexivity, symmetry and transitivity
Journal of the Experimental Analysis of Behavior, 1987, 47(3), 279-285

Mackay, H.A. (1985)

Stimulus Equivalence in rudimentary reading and spelling.

Analysis & Intervention in Developmental Disabilities, 1985, 5, 373-387

McManis, D.L. (1969)

Conservation and transitivity of weight and length by normals and retardates

Developmental Psychology, 1969, 1, 373-382

McManis, D.L. (1970)

Conservation, seriation and transitivity performance by retarded and average individuals.

American Journal of Mental Deficiency, 1970, 74, 784-791

Manabe, K. & Kawashima, T. (1993)

The role of vocal operant and auditory stimuli on stimulus equivalence in budgerigars.

Poster presented at the meeting for the Association for Behavior Analysis, Chicago, IL.

In Sidman, M. (1994)

Equivalence relations and behavior: A research story

Boston, MA: Authors Cooperative, Inc.

Piaget, J. (1928)

Judgement and reasoning in the child.

London: Routledge & Kegan Paul

Piaget, J. (1970)

Genetic epistemology.

Translated by Eleanor Duckworth.

New York: Columbia University Press, 1970

Rodewald, H.K. (1974)

Symbolic matching-to-sample by pigeons.

Psychological Reports, 1974, 34, 987-990

Romanes, G.J. (1888)

Animal Intelligence

In A.J. Riopelle (Ed.)

Animal Problem Solving (1967)

Middlesex, England; Penguin Books

Rumbaugh, D.M. (1977)

Language learning by a chimpanzee: The Lana Project.

New York: John Wiley

Saunders, R.R., Saunders, K.J. & Spradlin, J.E. (1990)

Long-term stability of equivalence relations in the absence of training or practice

American Journal on Mental Retardation, 1990, 95(3), 297-303

Saunders, K.J., Saunders, R.R., Williams, D.C., & Spradlin, J.E. (1993)

An interaction of instructions and training design on stimulus class formation:

Extending the analysis of equivalence.

The psychological Record, 43, 725-744

Savage-Rumbaugh, E.S. (1984)

Verbal behavior at a procedural level in the chimpanzee

Journal of the Experimental Analysis of Behavior, 1984, 41(2), 223-250

Savage-Rumbaugh, E.S., Rumbaugh, D.M., Smith, S.T. & Lawson, J (1980)

Reference: The linguistic essential.

Science, 210-922-925

Scheerer, M. (1963)

Problem-Solving

In A.J. Riopelle (Ed.)

Animal Problem Solving (1967)

Middlesex, England; Penguin Books

Schilmoeller K.J. & Etzel B.C. (1977)

An experimental analysis of criterion-related and noncriterion-related cues in "errorless" stimulus control procedures.

In B.C.Etzel, J.M.LeBlanc & D.M Baer (Eds.).

New developments in behavioural research : theory, method & application

Hillsdale NJ: Lawrence Erlbaum Associates

Schusterman, R.J. & Kastak, D. (1993)

A california sea-lion (*Zalophus californianus*) is capable of forming equivalence relations.

The Psychological Record, 1993, 43, 823-839.

Schusterman, R.J. & Krieger, K. (1984)

California sea lions are capable of semantic comprehension.

The Psychological Record, 34, 3-23

Sidman, M. (1971)

Reading and auditory-visual equivalences

Journal of Speech and Hearing Research, 1971, 14, 5-13

Sidman, M.(1986)

Functional analysis of emergent verbal classes.

In T. Thompson & M. D. Zeiler, (1986).

Analysis and integration of behavioural units

Hillsdale NJ: Lawrence Erlbaum Associates

Sidman, M.(1990)

Equivalence relations : where do they come from ?

In D.E. Blackman & H. Lejeune (Eds.)

Behaviour Analysis in theory and practice: contributions and controversies.

Hove, England: Erlbaum.

M. Sidman (1992)

Equivalence Relations : some basic considerations.

In S.C. Hayes & L.J. Hayes (eds.)

Understanding Verbal Relations

Reno, NV: Context Press

Sidman, M. (1994)

Equivalence relations and behavior: A research story

Boston, MA: Authors Cooperative, Inc.

Sidman, M. & Cresson, O. Jr. (1973)

Reading and cross-modal transfer of stimulus equivalences in severe retardation

American Journal of Mental Deficiency, 1973, 77(5), 515-523

Sidman, M., Cresson, O. Jr., & Willson-Morris, M. (1974)

Acquisition of matching to sample via mediated transfer

Journal of the Experimental Analysis of Behavior, 1974, 22(2), 261-273

Sidman, M., Kirk, B. & Willson-Morris, M. (1985)

Six-member stimulus classes generated by conditional-discrimination procedures

Journal of the experimental Analysis of Behavior, 1985, 43(1), 21-42

Sidman, M., Rauzin, R., Lazar, R., Cunningham, S., Tailby, W., & Carrigan, P. (1982)

A search for symmetry in the conditional discriminations of rhesus monkeys, baboons and children.

Journal of the Experimental Analysis of Behaviour, 1982, 37(1), 23-44

Sidman, M. & Tailby, W. (1982)

Conditional discrimination vs. matching to sample: an expansion of the testing paradigm

Journal of the Experimental Analysis of Behavior, 1982, 37(1), 5-22

Sidman, M. Wynne, C.K., Maguire R.W. & Barnes T. (1989)

Functional classes and equivalence relations

Journal of the Experimental Analysis of Behavior, 1989, 52(3), 261-274

Siemann, M.& Delius, J.D. (1993)

Implicit deductive reasoning in humans

Naturwissenschaften, 1993, 80, 364-366

Siemann, M. & Delius, J.D. (1994)

Processing of Hierarchic Stimulus Structures has Advantages in Humans and Animals.

Biological Cybernetics, 71, 531-536

Skinner, B.F. (1935)

The generic nature of the concepts of stimulus and response.

Journal of General Psychology, 12, 40-65

Skinner, B.F. (1938)

The behavior of organisms: An experimental analysis.

New York: Appleton-Century-Crofts

Skinner, B.F. (1957)

Verbal Behavior.

New York: Appleton-Century-Crofts

Smedslund, J. (1963)

Development of concrete transitivity of length in children.

Child Development, 1963, 34, 389-405

Smedslund, J. (1965)

The development of transitivity of length: A comment on Braine's reply.

Child Development, 1965, 36, 577-580

Spradlin, J.E., Cotter, V.W., & Baxley, N. (1973)

Establishing a conditional discrimination without direct training: A study of transfer with retarded adolescents

American Journal of Mental deficiency, 1973, 77(5), 556-566

Spradlin, J.E., & Dixon, M.H. (1976)

Establishing conditional discriminations without direct training: Stimulus classes and labels

American Journal of Mental Deficiency, 1976, 80(5), 555-561

Steele D. & Hayes S.C. (1991)

Stimulus Equivalence and arbitrarily applicable relational responding

Journal of the Experimental Analysis of Behavior, 1991, 56(3), 519-555

Sternberg, R.J. (1980)

Representation and process in linear syllogistic reasoning.

Journal of Experimental Psychology: General, 1980, 109, 119-159

Stromer, R. & Mackay, H.A. (1990)

A Note on the Study of Transitive Relations in Stimulus Sequences

Experimental Analysis of Human Behavior Bulletin, 1990, 8(1), 2-5

Stromer, R. & MacKay H.A. (1992)

Conditional stimulus control of children's sequence production.

Psychological Reports, 1992, 70, 903-912

Stromer, R. & Mackay, H.A. (1993)

Human sequential behavior: Relations among stimuli, class formation, and derived sequences

The Psychological Record, 1993, 43, 107-131

Stromer, R., Mackay, H.A., Cohen, M. & Stoddard, L.T. (1993)

Sequence learning in individuals with behavioural limitations

Journal of Intellectual Disability Research, 1993, 37, 243-261

Thorndike, E.L. (1898)

Animal intelligence; an experimental study of the associative processes in animals

In A.J. Riopelle (Ed.)

Animal Problem Solving (1967)

Middlesex, England; Penguin Books

Thayer, E.S., & Collyer, C.E. (1978)

The development of transitive inference : a review of recent approaches.

Psychological bulletin, 1978, 85(6), 1327-1343

Tomonaga, M., Matsuzawa, T., Fujita, K. & Yamamoto, J. (1991)

Emergence of symmetry in visual conditional discrimination by chimpanzees (Pan troglodytes)

Psychological Reports, 68, 51-60

VanBiervliet, A. (1977)

Establishing words and objects as functionally equivalent through manual sign training

American Journal of Mental Deficiency, 1977, 82(2), 178-186

Vaughan, W. (1988)

Formation of equivalence sets in pigeons

Journal of Experimental Psychology : Animal Behaviour Processes ,1988, 14(1), 36-42

Wright ,A.A., Cook, R.G., Rivera, J.J., Sands, S.F., Delius, J.D. (1988)

Concept learning by pigeons: matching to sample with trial unique video picture stimuli

Animal Learning and Behavior, 1988, 16(4), 436-444

Wulfert, E. & Hayes S.C. (1988)

Transfer of a conditional ordering response through conditional equivalence classes.

Journal of the Experimental Analysis of Behavior, 1988, 80, 125-144.

Youniss, J. & Murray, J.P. (1970)

Transitive inference with nontransitive solutions controlled.

Developmental Psychology, 1970, 2, 169-175

APPENDIX I

Case Study of Stimulus Equivalence Performance by an Autistic Adult

This experiment was an investigation of stimulus equivalence performance by an adult male (AS) diagnosed as autistic. Eikeseth and Smith (1992) had carried out an investigation of stimulus equivalence performance with four high-functioning autistic children. They suggested that autistic subjects may be an appropriate population with which to examine the conditions under which equivalence relations develop, as they seem "less likely to display untrained relations (such as equivalence relations) than other human populations" (Eikeseth and Smith, 1992, p.124). Subject AS also had a specific educational history that suggested that it might be instructive to repeat aspects of experiment 1 with him as a subject.

Experiment 1 examined the role of naming in equivalence performance using upper and lower case Greek letters and their printed names as stimuli. Subject AS had received teaching in ancient Greek language over a period of several years. Thus, he had a unique history of familiarity with the items used as training and test stimuli, but some data from clinical tests suggested that AS might have difficulty with tasks such as stimulus equivalence.

One of the major debates in equivalence research had been the role of naming and language in equivalence class formation (see chapter 5). Having had this extensive history of instruction in Greek, it seemed likely that AS would have learned names for the stimuli that were used for training and testing in experiment 1. If AS was given conditional discrimination training using these stimuli, would he spontaneously use these names, and would there be any evidence that they facilitated appropriate equivalence performance, or not, as the case may be.

There was no guarantee that AS would recognise the stimuli as Greek letters if he was not explicitly told that that was what they were. This seems to relate to the relevance of the property of reflexivity in equivalence research. Reflexivity is often assumed in equivalence research and not tested, even when the properties of symmetry and transitivity are assessed. Yet reflexivity is vital as it fixes the identity of a stimulus across situations. In the context of equivalence testing this generally means that the subject will treat a stimulus as the same stimulus whether it functions as a sample or as a comparison. For instance, a subject could not demonstrate symmetry of a taught A1-B1 relation if that subject treated A1 as a

different stimulus when it served as a comparison rather than a sample. Thus, if AS did not recognise the stimuli as Greek letters unless they appeared in the context of a Greek lesson, then there is no reason to assume that there would be any stimulus names available to him during testing.

The aim of this experiment then, was to give subject AS conditional discrimination training using upper and lower case Greek letters and their printed names as stimuli, and then test to see if AS had formed appropriate equivalence classes. In conjunction with this, AS was given a number of naming tests to see if he spontaneously used any names (Greek or other) for the stimuli, and if there was any evidence that these names influenced his performance on stimulus equivalence tests.

Method

Subject

The subject in this experiment was AS, a 25 year old male. It had previously been suggested that a diagnosis of autism or Asperger's syndrome would be appropriate for AS. AS had received quite a high level of educational input when younger. He lived semi-independently in a housing scheme in a town near St. Andrews.

Prior to testing in this experiment, AS had recently been assessed on a number of clinical measures by a qualified clinical psychologist; the Wechsler Adult Intelligence Scale-Revised (WAIS-R), the Wechsler Memory Test-Revised (WMS-R), the Rey Osterreith Figure Test-Copy and Recall, the Trail Making Test and the Word Fluency Test.

AS's performance on the WAIS-R placed him in the borderline range of intelligence, with his verbal score slightly higher than his performance score. He performed best on tests of information, vocabulary and similarities, and was relatively poor at timed tests requiring manipulation of numbers, figures and symbols. He also showed problems with attention and sustained concentration, mental calculation, visuo-motor speed and forward planning.

Performance on the WMS-R on tests of verbal memory and learning were much as would be predicted from his intellectual level and any problems seemed more likely to be due to attentional difficulties rather than specific memory deficits.

On the Rey Osterreith figure copying and figure recall tasks AS's final performance was fairly complete and accurate but he seemed to have trouble with organisation of the task. The Trail Making Test assesses visual tracking, attention and mental sequencing and AS's performance was poor, suggesting problems with concentration and organisational skills, and impaired mental sequencing and visuo-motor speed. Finally, AS showed poor performance on the Word Fluency Test of spontaneous word production. This was somewhat surprising given his relatively good verbal performance on the WAIS-R, suggesting he has problems with tasks which lack structure and involve initiative and spontaneity.

In general AS appeared to be a somewhat anxious individual, and tended to set very strict time limits on test sessions, often restricting the amount of testing possible following baseline training. He was also easily distracted and showed a fairly limited concentration span, which also limited the amount of testing possible.

Stimuli

The stimuli used for training and testing were the Greek letters Gamma and Delta shown in Table A:1. These were presented on photographic cards approx. 13 cm's x 9 cm's, covered with clear plastic, one stimulus to each card. The stimuli appeared as black forms on a white background.

These stimuli were also used for one of the naming tests given to subject AS during the course of the experiment. Two other sets of stimuli were used for two different naming tests at different times. One was a set of eight Greek letters - their upper and lower case forms and their printed names (see Table A:2). This was the same naming test given to the subjects in experiment 1 before and after testing. The third test was the upper and lower case forms of all the letters of the Greek alphabet (see Table A:3).

Procedure

The basic procedure was that before testing AS would be given one of the three naming tests to check for any spontaneous naming of the stimuli. AS would then be given match-to-sample training on the conditional relations A1-B1, A2-B2, B1-C1 and B2-C2. AS was then tested on the A-B and B-C mixes. If AS reached criterion on these relations he was then tested for the emergence of C-A equivalence relations. The naming test would then be repeated to see if AS had either produced more names for the stimuli during the course of testing, or showed any evidence of producing names consistently for any of the stimuli.

Slightly different instructions were given according to which naming test was used:

1. The first naming test consisted of the two classes of A, B and C stimuli used in training. In this test AS was shown the stimuli individually and asked if he thought that form had a name, or if it should be called anything. AS was not explicitly told the stimuli were Greek letters.
2. The second naming test consisted of the eight classes of A, B and C stimuli as used in experiment 1. Two of these classes were the classes of stimuli used in training. Once again AS was not explicitly told that these stimuli were Greek letters. Rather, he was again asked if he thought any of the stimuli had a name or should be called anything. The aim of this test was to see if AS would be more likely to name the stimuli used in testing, or produce consistent names for these stimuli, than for similar stimuli not used in training.
3. The third naming test consisted of the upper and lower case forms of all the letters in the Greek alphabet. Before being given this test, AS was explicitly told that these stimuli were Greek letters, and he was asked to try and name them. This was to see if AS was aware of the names of the stimuli and would be more likely to produce names for these stimuli if explicitly told that the stimuli were Greek letters.

During conditional discrimination training, a trial and error training procedure was used. Initially the sample stimulus was placed in front of AS and he was asked to look at it and touch it; this was to establish an observing response. Having touched the sample, AS was then shown the two comparison stimuli which were placed side by side underneath the sample stimulus. AS was asked to select the comparison stimulus that "went with" the sample stimulus. It proved difficult to establish a consistent observing response with AS, and so, after the first few sessions, this procedure was changed slightly. When the sample stimulus was placed in front of AS, it was rotated 90° to the left or right, or was upside down. AS was required to turn the stimulus back to the correct position to try and ensure that he observed the sample properly. The comparison stimuli were presented as previously.

On the conditional discrimination training trials, criterion was set at five consecutive correct responses. Once criterion had been reached on A1-B1 and A2-B2, these tasks were combined to form the A-B mix, and similarly, when criterion was reached on B1-C1 and B2-C2, these tasks were combined to form the B-C mix. On these tasks the sample stimulus varied between A1 and A2 or between B1

and B2, each stimulus appearing five times in a block of ten trials. The left/right position of the comparison stimuli varied randomly on all training trials.

During training and testing, the stated experimental procedure was sometimes constrained by a tendency on the part of AS to set strict time limits on the test sessions. This occasionally meant that the full range of tests could not be given on one day, or that it was not possible to complete all of the post test naming test.

Results

AS was given training and testing on eight separate occasions.

On the first session AS reached criterion on both A1-B1 and A2-B2 fairly easily, although performance on the A-B mix was poor (see Fig A:1). AS also reached criterion on the two B-C relations, although this training took longer and performance was not as stable. However, performance on the B-C mix was slightly better than on the A-B mix. AS was tested on the C-A equivalence test but gave no evidence of having derived equivalence relations (C-A equivalence - 40%). During this session all AS's responses were somewhat unclear and no tests for naming were given.

In session 2, prior to training, AS was tested for naming of the eight Greek letters (upper and lower case figures, and printed names). AS did not give appropriate Greek names for any of the upper case letters and of the lower case letters only named theta correctly. He did however give names to most of the stimuli, generally based on their similarity to standard letters. AS pronounced the names of six of the printed words correctly, and also spelled out the letters for xi and mispronounced phi. It was not possible to give a post-training naming test as AS terminated the session immediately after baseline discrimination training and testing.

Performance on the conditional discriminations varied considerably. Performance on A1-B1 was poor and did not reach criterion, yet on A2-B2 performance was perfect. A1-B1 was retrained and performance was again poor, suggesting responding was actually based on a preference for stimulus B2. Performance on the B-C relations was also poor, never reaching criterion.

In session 3 AS was given the naming test for the eight Greek letters before and after training. AS gave names for most of the stimuli in these tests. AS named all the printed names correctly both times. He was also more likely to give correct Greek letter names for the lower case stimuli, and also to use these names consistently, than for the upper case stimuli. (See table A:4) On the conditional discriminations, AS reached criterion on both A1-B1 and A2-B2, although his performance took some time to stabilise (see Fig. A:3). When tested on the A-B mix his performance was at chance level.

In session 4 AS was given the third naming test, which presented all the upper and lower case letters of the Greek alphabet. On this task AS produced names for 20 of the upper case stimuli (total 24), 9 of which were correct. AS produced names for 23 of the lower case stimuli, 15 of which were correct. Performance on the B-C conditional discrimination was very good and AS also responded at 100% when tested on the B-C mix (see Fig. A:4). After conditional discrimination training AS was shown the six stimulus cards for the A, B and C training stimuli. He gave appropriate names for the lower case (B) stimuli, and the printed name (C) stimuli. However he did not give appropriate names for either of the A stimuli.

In session 5, AS was given the naming test consisting of the A, B and C parts of the eight Greek letters. He gave names for six of the upper case stimuli, two of which were correct although neither were the test stimuli. He gave names for all eight lower case stimuli, six of which were correct, including both test stimuli, and gave correct names for all the printed name stimuli. AS was given training on the A1-B1 and A2-B2 conditional discriminations (see Fig. A:5). Performance on these was poor and neither discrimination reached criterion. AS was then given training on B1-C1 to see how far this would compare to the A-B performance, and immediately reached criterion.

In session 6, no naming tests were given. AS was given training on both A-B and B-C discriminations (see Fig. A:6). On this occasion AS reached criterion on all four discriminations, although his performance took some time to stabilise on A2-B2.

In session 7, again no naming tests were given, and AS received training on all four conditional discriminations (see Fig. A:7). AS reached criterion on both A-B discriminations, although again, performance on A2-B2 was poor. When trained on the B-C discriminations, performance reached criterion immediately. The A2-B2 discrimination was re-trained, and this time reached criterion immediately. AS was

then tested for B-A symmetry, one of the components of equivalence relations. However performance on this test was at chance level.

Training and testing for session 8 took place over two consecutive days to try and allow a more thorough examination of AS's performance on the equivalence test. Before and after testing, AS was given the third naming test, consisting of the upper and lower case letters of the Greek alphabet. In both naming tests AS gave names for most of the stimuli. However he was slightly more likely to give names for the lower case stimuli, and it was more likely that these would be the appropriate Greek letter names. It was also slightly more likely that he would produce consistent names for the lower case stimuli (see Table A:5). Of the uppercase test stimuli, AS only named the gamma (A1) stimulus, and this naming was not consistent from pre- to post- test. AS failed to name the delta (A2) stimulus at all. However AS correctly named both the lower case (B1 and B2) stimuli in both the naming tests.

AS was trained on both the A1-B1 and A2-B2 conditional discriminations (see Fig. 8:8). He reached criterion on both these tasks, although performance was not particularly good. In an attempt to maintain stable performance in preparation for equivalence testing, AS was tested on the AB mix, but this time with feedback on all trials. This seemed to help maintain performance on a task where the sample stimuli changed randomly (A-B mix - 80%). AS was then trained on the B-C discriminations, and performance was highly accurate. AS was then tested on both the B-C and A-B mixes, without reinforcement. Performance on the B-C mix was perfect, while performance on the A-B mix had deteriorated slightly from when all trials were reinforced (A-B mix - 62%). AS was given the C-A equivalence test, but performance was close enough to chance level to suggest that AS had not derived equivalence classes (C-A equivalence - 60%).

Discussion

The tests given on a number of occasions during this sequence of training and testing give no evidence that AS had derived any of the relations of stimulus equivalence. AS was directly tested on the C-A equivalence test on two occasions, following the first training session, and following the eighth (last) training session. On both occasions, performance was around chance level. On one occasion AS was tested for the emergence of B-A symmetry, one of the necessary properties of stimulus equivalence. Performance on this relation was also at chance level, even

though AS had just been trained on the A-B relations, and had reached criterion on both these discriminations. This suggests that at no time did AS derive any of the properties of stimulus equivalence as a result of the conditional discrimination training given.

It generally proved much harder to establish stable performance on the A-B relations (upper case letter - lower case letter) than on the B-C relations (lower case letter - printed name). Given this difficulty in establishing stable baseline performance, the failure of AS to demonstrate equivalence relations is understandable. It might have been interesting to test AS for the emergence of C-B symmetry. Performance on the trained B-C relations was generally highly accurate. This may have made it more likely that AS would display the C-B symmetrical counterpart of these relations. This might have given some indication of whether AS would derive *any* of the properties of equivalence relations. However, when AS was tested on B-A symmetry, he had just displayed criterion performance on both the A-B relations, which would suggest that AS genuinely failed to display any equivalence relations, rather than his failure on the symmetry test being due to unstable performance on the baseline relations.

One of the reasons for interest in performance by AS on this equivalence task was to see whether his experience with Greek letters would influence his performance on these tests. There is some evidence to suggest that his performance on the conditional discriminations, and familiarity with the stimuli are linked. In all the naming tests AS was more likely to give correct names for the lower case stimuli than for the upper case stimuli. He was also more likely to give the same names to the lower case stimuli both before and after conditional discrimination training, suggesting that these names would have been available to him during discrimination training. When the upper case stimuli were tested for names AS never gave appropriate Greek letter names for the test stimuli. However, he generally gave appropriate names for the lower case test stimuli. If AS had derived B-A symmetry, then it is possible that these appropriate lower case names would have transferred to the upper case stimuli. The failure of AS to demonstrate B-A symmetry is consistent with the failure of these names to transfer to the A stimuli. Similarly, AS performed better on the B-C discriminations than on the A-B discriminations, which also fits with his superior naming of the lower case (B) stimuli. AS was generally able to name the printed name (C) stimuli, although it is not clear whether he was reading the word as he was shown it, or whether he was already familiar with that name and knew that it was also the name of a Greek letter.

As subject AS was already able to name the B stimuli, and at the very least could read the names of the C stimuli, it is possible that in effect, AS "knew" the B-C relation prior to the conditional discrimination training. This would be one explanation why this relation seemed to be established so much more easily than the A-B conditional discrimination.

While there seems to be some connection between performance by AS on the conditional discriminations and his ability to name the stimuli, it is not clear that this is a causal relationship. If naming does cause equivalence, as several researchers have suggested, then it would be likely that AS should have shown some improvement in equivalence performance over the course of testing. After all, AS could fairly reliably name the B and C test stimuli. AS was explicitly taught, on a number of occasions, to link the A stimuli with the B stimuli. If naming does cause equivalence, then some improvement in A-B performance over the course of testing would be expected. Yet there was no evidence of this, and even when the A-B relations were at criterion, naming of the B stimuli did not facilitate the development of B-A symmetry.

Of course, while the naming tests were designed to check if AS was *able* to name the stimuli, they cannot check if AS then *used* these names during conditional discrimination training and testing. During training and testing AS very rarely overtly labelled the stimuli. While he may have been covertly labelling the stimuli, he might also not have been using the stimulus names at all. Possibly if AS had named the stimuli during training and testing, this would have resulted in equivalence responding. However, the evidence from this experiment would also seem to support Sidman's suggestion that naming may result from equivalence relations, rather than the other way round.

There is no clear evidence that the context of the naming test caused a difference in the way AS named the stimuli. In the test with eight Greek letters where AS was not told the stimuli *were* Greek letters, AS spontaneously produced several Greek letter names, although not always appropriately. In the third naming test when AS was explicitly asked to give the stimuli their Greek letter names, he was slightly more inclined to give Greek names to all the stimuli, but he still occasionally gave other stimulus names, e.g. stimulus A2 which was upper case delta (Δ) was labelled "triangle".

It would appear then that AS's previous experience with Greek letters did not facilitate the development of equivalence relations.

Table A:1 Training and Test Stimuli

Stimulus Class	A	B	C
1	Γ	γ	gamma
2	Δ	δ	delta

Table A:2 Stimuli for Second Naming Test

Stimulus class	A	B	C
1	Γ	γ	gamma
2	Δ	δ	delta
3	Θ	θ	theta
4	Λ	λ	lambda
5	Ξ	ξ	xi
6	Φ	ϕ	phi
7	Σ	σ	sigma
8	Ω	ω	omega

Table A:3 Stimuli for Third Naming test

Stimulus Name	Upper Case Character	Lower Case Character
Alpha	A	α
Beta	B	β
Gamma	Γ	γ
Delta	Δ	δ
Epsilon	E	ϵ
Zeta	Z	ζ
Eta	H	η
Theta	Θ	θ
Iota	I	ι
Kappa	K	κ
Lambda	Λ	λ
Mu	M	μ
Nu	N	ν
Xi	Ξ	ξ
Omicron	O	\omicron
Pi	Π	π
Rho	P	ρ
Sigma	Σ	σ
Tau	T	τ
Upsilon	Y	υ
Phi	Φ	ϕ
Chi	X	χ
Psi	Ψ	ψ
Omega	Ω	ω

Table A:4 Session 3 - Results of Pre- and Post Test Naming Tests

	Pre-Test Names	Post-Test Names	Consistent Names
Upper Case Letters	7 names (2 correct, not test)	8 names (2 correct, not test)	3 names (2 correct)
Lower case Letters	8 names (5 correct, both test)	8 names (4 correct, both test)	6 names (4 correct, both test)
Printed Names	8 names (all correct)	8 names (all correct)	8 names (all correct)

The cells give the total number of names given to the stimuli. The figures in brackets indicate the number of these names that were appropriate Greek letter names, and if these included the test stimuli.

Table A:5 Session 8 Results of Pre- and Post- Test naming Tests

	Pre-Test Names	Post-Test Names	Consistent Names
Upper Case Letters	19	19	16
	(11 correct, not test)	(11 correct, not test)	(11 correct, not test)
Lower Case Letters	23	23	19
	(18 correct, both test)	(16 correct, both test)	(16 correct, both test)

The cells give the total number of names given to the stimuli. The figures in brackets indicate the number of these names that were appropriate Greek letter names, and if these included the test stimuli.

Fig A:1 Session 1 Results

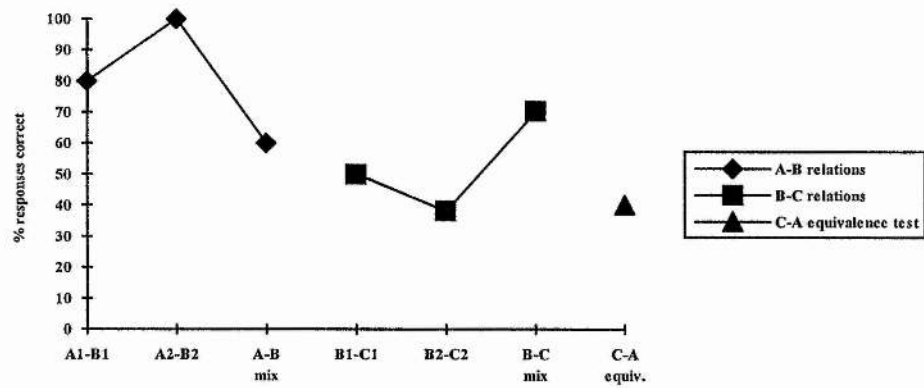


Fig A:2 Session 2 Results

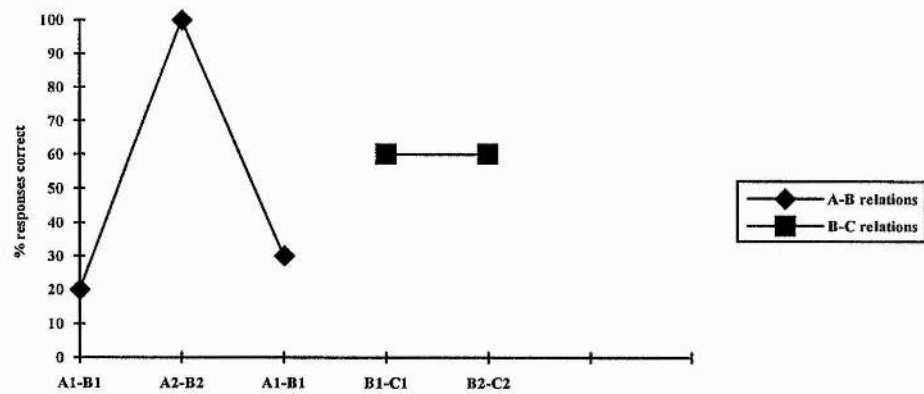


Fig A:3 Session 3 Results

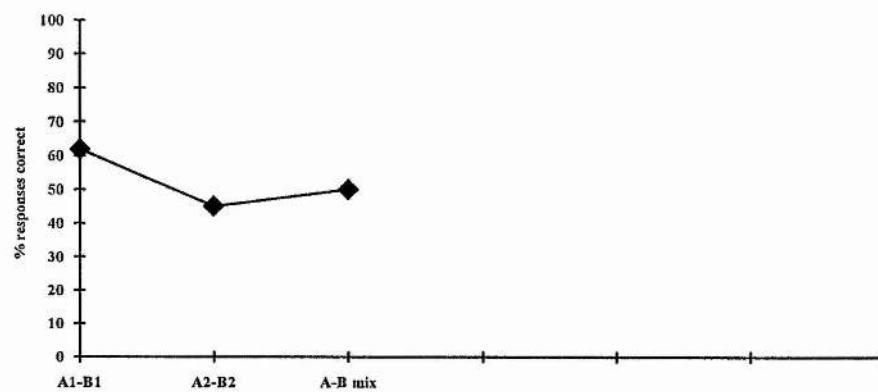


Fig A:4 Session 4 Results

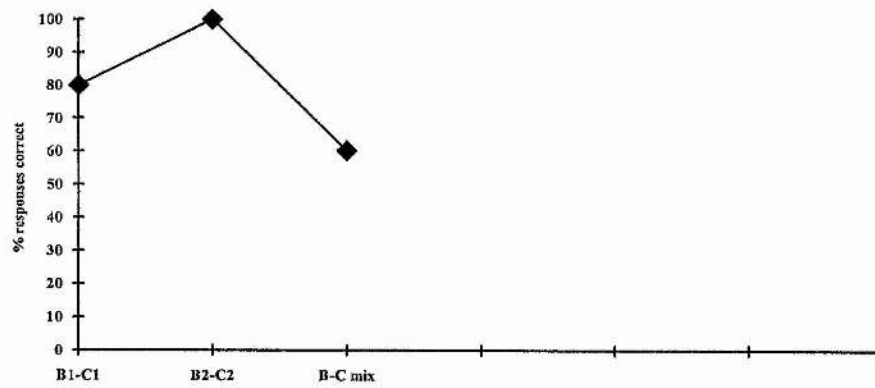


Fig A:5 Session 5 Results

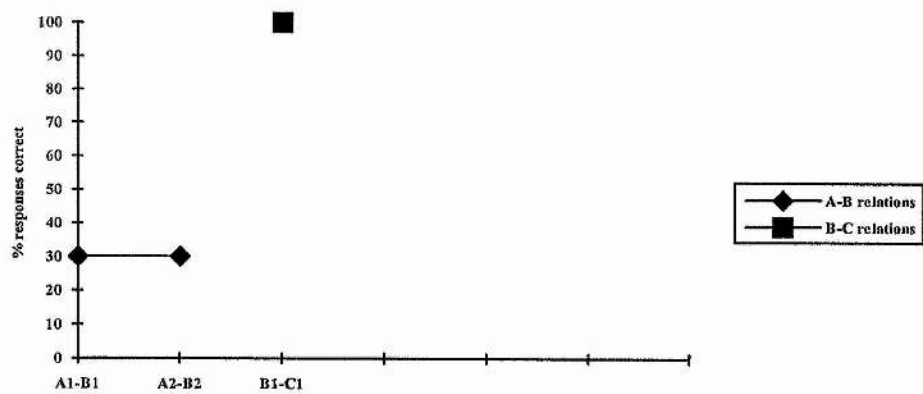


Fig A:6 Session 6 Results

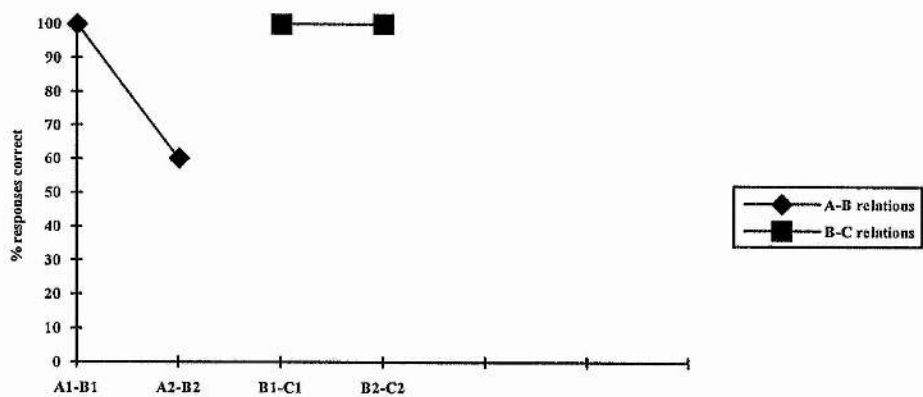


Fig A:7 Session 7 Results

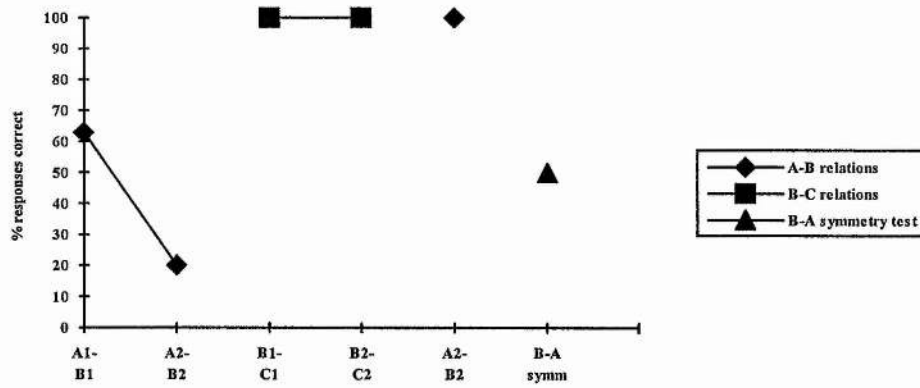


Fig A:8 Session 8 Results

